



UNIVERSIDAD CARLOS III DE MADRID

working
papers

Working Paper 11-18 (13)
Statistics and Econometrics Series
June 2011

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

DENSIDAD DE PREDICCIÓN BASADA EN MOMENTOS CONDICIONADOS Y MÁXIMA ENTROPÍA. APLICACIÓN A LA PREDICCIÓN DE POTENCIA EÓLICA

Miguel Ángel Bermejo*, Daniel Peña** e Ismael Sánchez*

Resumen

El cálculo de predicciones puntuales junto con su incertidumbre en forma de intervalo es, en la mayoría de aplicaciones, insuficiente. Especialmente cuando estemos asumiendo no linealidad en los datos, puesto que en estos casos, podrían existir incluso cambios en la distribución. Por ello será necesario, además de la predicción puntual, obtener una estimación de la densidad condicionada de la variable en el futuro dado su comportamiento actual, es decir, la densidad predictiva. En este trabajo proponemos una estimación de la densidad predictiva empleando diferentes distribuciones paramétricas como son la Normal Truncada, la Normal Censurada, la Beta y la de Máxima Entropía. Dichas distribuciones serán calculadas empleando los momentos condicionados estimados mediante un método de estimación recursiva. Se aplica el procedimiento a datos provenientes de energía eólica.

Palabras clave: Densidad predictiva; Máxima entropía; Momentos condicionados; Estimación recursiva; Potencia eólica

* Department of Statistics, Universidad Carlos III de Madrid, Av. Universidad 30, 28911 Leganés (Madrid), Spain, e-mail: miguelangel.bermejo@uc3m.es and ismael.sanchez@uc3m.es.

** Department of Statistics, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain e-mail: daniel.pena@uc3m.es

Acknowledgements: Work supported by Ministerio de Educación y Ciencia, project SEJ2004-03303 and European SAFEWIND-EU project 213740.

Densidad predictiva basada en momentos condicionados y máxima entropía. Aplicación a la predicción de potencia eólica

Miguel Ángel Bermejo, Daniel Peña e Ismael Sánchez

28 de junio de 2011

Resumen

El cálculo de predicciones puntuales junto con su incertidumbre en forma de intervalo es, en la mayoría de aplicaciones, insuficiente, especialmente cuando estamos asumiendo no linealidad en los datos, puesto que en estos casos, podrían existir incluso cambios en la distribución. Por ello será necesario, además de la predicción puntual, obtener una estimación de la densidad condicionada de la variable en el futuro dado su comportamiento actual. A esta distribución condicionada se le conoce con el nombre de densidad predictiva. En este trabajo proponemos una estimación de la densidad predictiva empleando diferentes distribuciones paramétricas como son la Normal Truncada, la Normal Censurada, la Beta y la de Máxima Entropía. Dichas distribuciones serán estimadas de manera adaptativa empleando los momentos condicionados estimados mediante un método de estimación recursiva. Se aplica el procedimiento a datos provenientes de energía eólica.

Palabras clave: Densidad predictiva; Estimación recursiva; Máxima entropía; Potencia eólica.

1. Introducción

La desventaja de la energía eólica en el mercado eléctrico respecto a otras fuentes de energía es su incertidumbre. Para minimizar esta incertidumbre es necesario obtener predicciones que permitan su integración en el mercado eléctrico. Estas predicciones deben ser a muy corto plazo, debido a la existencia de un mercado eléctrico intradiario.

La realización de predicciones puntuales de potencia eólica empleando modelos estadísticos ha sido ampliamente estudiada en años recientes (Sánchez, 2006; Costa et al., 2008; Giebel et al., 2011). El

problema es que las predicciones puntuales no son suficientes, ya que en ocasiones se hace necesario obtener la estimación de la densidad predictiva, para así tener un mejor conocimiento acerca del comportamiento de las predicciones (Diebold y López, 1996). En los últimos años han sido múltiples las propuestas de métodos de estimación de densidades predictivas, como puede verse en Tay y Wallis (2000), Mitchell (2008) o Clements y Harvey (2010).

En particular, la utilización de funciones de densidad predictiva para la energía eólica permitirá, entre otras cosas, mejorar la gestión de las reservas de energía eléctrica, mejorar la programación de las centrales de energía eléctrica convencionales y minimizar el riesgo que cometen los agentes que ofrecen energía eólica en el mercado eléctrico.

El problema de estimación de densidades predictivas no ha sido muy tratado cuando trabajamos con la potencia eólica. En este caso la estimación es más compleja aún que en los casos tratados tradicionalmente, ya que la variable es no gaussiana, y además está acotada entre 0 y la potencia nominal. Como en general se trabaja estandarizando la potencia producida por la potencia nominal, los datos estarán acotados entre 0 y 1.

Entre las propuestas realizadas en la literatura, existen diferentes puntos de vista desde los que abordar el problema de la estimación de la densidad predictiva. Una posible opción es trabajar directamente con la distribución de la velocidad del viento para luego emplear la relación no lineal existente entre velocidad y potencia generada para poder estimar la distribución de la potencia (p.ej., Lange, 2005; Carta et al. 2009). Otra propuesta similar, aunque más compleja, se basa en la realización de múltiples predicciones de variable meteorológicas a partir de las cuales se estima la distribución de la producción (p.ej., Nielsen et al. 2004; Pinson y Madsen 2009; Taylor et al. 2009).

Por último otras propuestas diferentes se basan la estimación de la densidad predictiva a partir de los datos históricos de la potencia y de sus predicciones (p.ej., Bludszuweit et al., 2008; Al-Awami y El-Sharkawi, 2009; Lau y McSharry, 2010; Pinson, 2010). Este punto de vista será el empleado en nuestra propuesta de estimación. Empleando las predicciones de potencia estaremos de manera adaptativa los momentos condicionados de los errores de predicción. Con estos momentos condicionados, estaremos la senda de predicciones predictiva ajustando diferentes distribuciones paramétricas.

2. Energía eólica

La estimación de la densidad predictiva dependerá de los datos con los que se está trabajando. Por ejemplo, si los residuos del modelo empleado son gaussianos, el problema de estimar la densidad predictiva quedará reducido a la estimación de la media y la varianza condicionales. Sin embargo, esto no será adecuado en otras ocasiones en las que, por ejemplo, se pueden encontrar densidades predictivas no gaussianas. El objetivo de este capítulo es mejorar la estimación de la densidad predictiva de las predicciones puntuales de potencia eólica. Es decir, para unas predicciones puntuales dadas queremos estimar cuál es su densidad condicionada.

Para analizar los diferentes métodos propuestos a lo largo del capítulo vamos a emplear datos de potencia eólica media producida horariamente en 6 parques eólicos diferentes. Para cada uno de dichos parques eólicos tendremos la producción media horaria observada y las predicciones de dicha producción a diferentes horizontes. Desconocemos cómo han sido calculadas las predicciones, por lo que no tenemos ninguna información previa acerca de la distribución de los datos. La frecuencia de dichos datos es horaria.

2.1. Análisis descriptivo de la distribución de la potencia generada

La principal característica de la potencia eólica es que es una variable acotada, entre 0 y la potencia nominal, es decir, el valor máximo que puede producir un aerogenerador o un parque eólico, dependiendo de los datos con los que estemos trabajando. En general, se trabaja estandarizando la potencia producida por la potencia nominal, con lo que los datos estarán acotados entre 0 y 1. Los datos que hemos empleado en este capítulo pertenecen a la potencia eólica horaria generada por diferentes parques eólicos. Para esas potencias medidas tenemos las predicciones puntuales elaboradas por cierto agente para diferentes horizontes de predicción. El objetivo es estimar la función de densidad predictiva a partir de la serie temporal de predicciones y los datos históricos observados, sin tener ninguna información acerca del método empleado en la elaboración de las predicciones. Como vemos en la figura <ref>Fig-Potencia</ref> la potencia eólica es una variable no estacionaria que puede cambiar de manera abrupta. En la figura pequeña en la que mostramos un tramo de la serie vemos

cómo la serie alterna rápidos cambios de nivel.

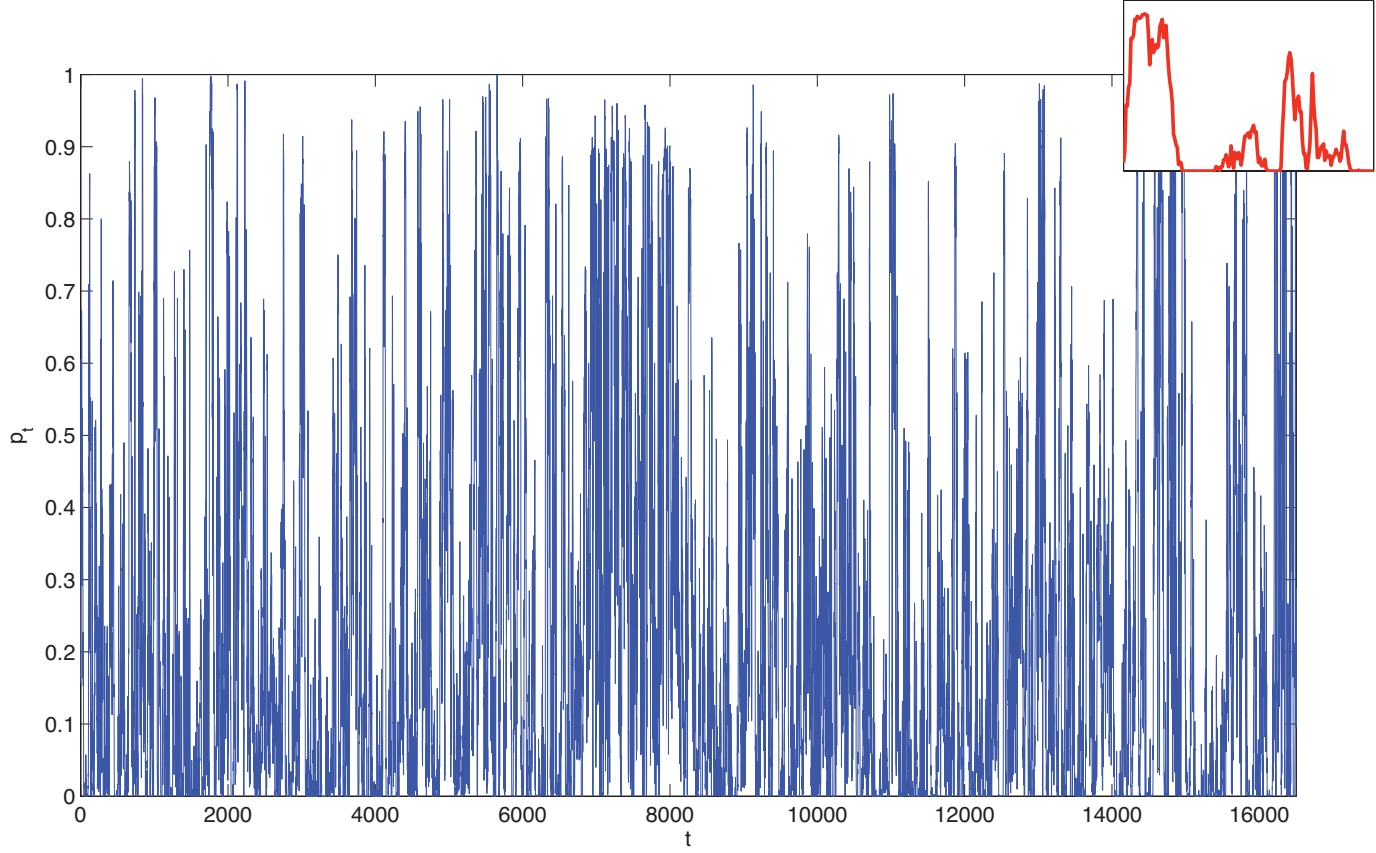


Figura 1: Potencia producida por un parque eólico, datos horarios

El objetivo de este trabajo es proponer un procedimiento para estimar la densidad predictiva de las predicciones para diferentes horizontes de predicción. En la figura 2 vemos los gráficos de caja de la potencia medida para diferentes niveles de las predicciones. En dichas figuras vemos cómo existen asimetrías y varianzas que dependen del nivel de predicción independientemente del horizonte. Es decir, los errores de predicción tienen algunas características que pueden observarse a cualquier horizonte de predicción.

2.2. Notación

Para facilitar la lectura, vamos a introducir la notación empleada a lo largo del mismo. Dado el carácter estocástico del viento, la potencia eólica generada durante un periodo de tiempo puede interpretarse como un proceso estocástico, cuyos valores observados serán denotados por p_t . Por simplicidad

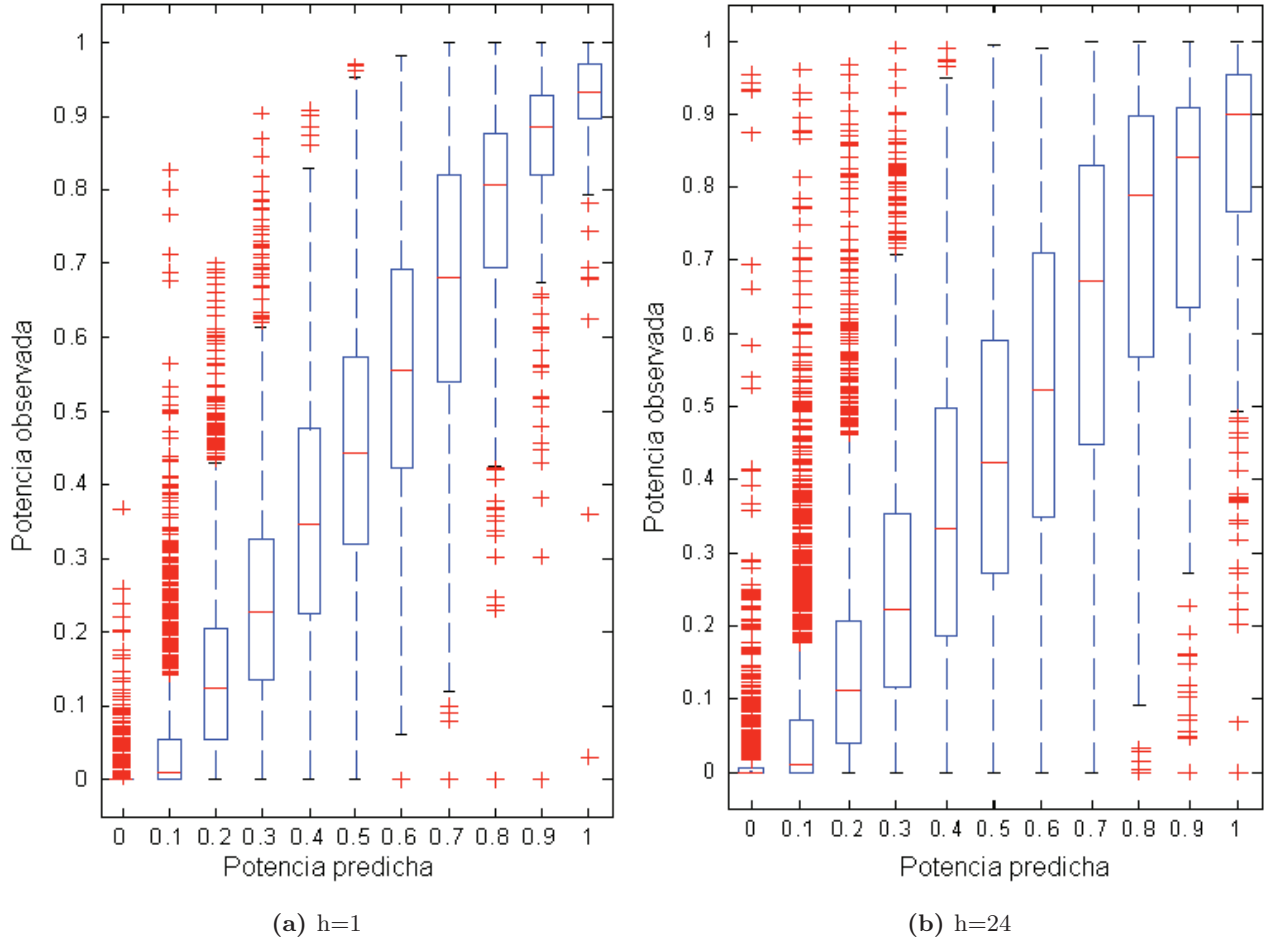


Figura 2: Gráfico de caja para de la potencia medida respecto a diferentes niveles de la predicción puntual

en la notación siempre usaremos p_t , tanto para referirnos al proceso como a los valores observados. El objetivo del trabajo será caracterizar a la variable aleatoria p_{t+h} condicionada a la información en el instante de tiempo t . Para caracterizar dicha variable aleatoria p_{t+h} tenemos que conocer su función de distribución F_{t+h} o su función de densidad f_{t+h} . Denotaremos a la predicción de la variable aleatoria p_{t+h} a horizonte h como $\hat{p}_{t+h|t}$. El error de predicción a horizonte h será $e_{t+h|t} = p_{t+h} - \hat{p}_{t+h|t}$, y su momento condicionado de orden 1 será $m_{t+h|t}^{(0)} = E_t[e_{t+h|t}]$. Los momentos centrales del error de predicción $e_{t+h|t}$ serán $m_{t+h|t}^{(j)} = E_t\left[\left(e_{t+h|t} - m_{t+h|t}^{(0)}\right)^j\right]$, donde $j = 2, 3, \dots$

3. Densidad predictiva en energía eólica

La estimación de la densidad predictiva de la potencia eólica es un área de investigación que se ha empezado a estudiar en años recientes. Existen varios puntos de vista desde los que abordar el problema.

3.0.1. Empleando la distribución de la velocidad del viento

El principal componente de la potencia eólica producida por un parque es la velocidad del viento, por lo tanto parece evidente que el estudio de la velocidad del viento será una herramienta adecuada para la modelización de la potencia generada a través de predicciones de la velocidad. En el trabajo de Carta et al. (2009) se hace una gran revisión sobre diferentes estimaciones de distribuciones para la velocidad del viento.

Entonces, una vez estimada la distribución de la velocidad del viento, la predicción probabilística de la potencia eólica se construye a partir de la curva de potencia que mide la relación existente entre velocidad y potencia (p.ej., Lange 2005). El principal problema es que dicha relación es no lineal y cambiará dependiendo del tiempo, del aerogenerador, de la temperatura ambiental, de la dirección del viento, etc.

3.0.2. Empleando "*ensemble predictions*"

En meteorología es habitual suministrar la predicción de diferentes variables ambientales para un periodo de tiempo en forma de un conjunto de trayectorias posibles a partir de diferentes condiciones iniciales y empleando diferentes modelos (Epstein, 1969). A ese conjunto de trayectorias se le conoce como "*ensemble predictions*". Entonces, una vez generadas las predicciones *ensemble* la densidad predictiva de la producción de energía eólica se calcula empleando las curvas de potencia que relacionan las variables meteorológicas empleadas con la potencia eólica (Nielsen et al. 2004, Pinson y Madsen 2009, Taylor et al. 2009). En este caso, además del problema de la relación no lineal existente comentado en el anterior punto, se le añade la dificultad de obtener las predicciones *ensemble*, ya que tienen que ser suministradas por una agencia meteorológica.

3.0.3. Empleando regresión cuantílica

Estos métodos consisten en aplicar una regresión por cuantiles sobre la producción eólica en función de un conjunto de variables explicativas sin tener que asumir ningún tipo de estructura en la densidad condicionada de la potencia. Bremnes (2006) comprobó su eficiencia comparándolo con un modelo Gaussiano y un estimador de Nadaraya-Watson para la función de distribución, obteniendo resultados muy parejos para todos los métodos, dando ventaja al método de Nadaraya-Watson por cuestiones de sencillez en la computación. En Moller et al. (2008) se propone un procedimiento adaptativo para la regresión por cuantiles que mejora el tiempo de computación.

3.0.4. Empleando la predicción de potencia eólica

Otra de las opciones existentes consiste en ajustar una distribución de probabilidad condicionada de la potencia a partir de los datos históricos de la potencia medida y de predicciones de la misma. La principal dificultad está en que la potencia eólica es una variable con soporte acotado y no gaussiana, además de ser una variable con un comportamiento no estacionario. Por ello, Lau y McSharry (2010) y Pinson (2010) proponen realizar una transformación logística, con el proposito de conseguir una distribución Normal, empleando a continuación diferentes modelos paramétricos para realizar las predicciones puntuales y estimar la densidad predictiva empleando las funciones de distribución Normal Truncada y Normal Censurada.

Otra opción es ajustar una distribución paramétrica condicionada a partir de la estimación de momentos condicionados de la potencia. Bludszuweit et al. (2008) y Al-Awami y El-Sharkawi (2009) lo hicieron empleando distribuciones como la Beta y la distribución generalizada del valor extremo, a partir de los dos primeros momentos.

Este es el camino que hemos seguido para nuestra propuesta de estimación de la densidad predictiva de la potencia eólica. Para ello proponemos emplear un método adaptativo que nos permite estimar los momentos centrales condicionados de los errores de predicción mediante modelos de regresión adaptativa. Las variables explicativas de estas regresiones serán las predicciones puntuales de la potencia. Una vez obtenidos los momentos, los emplearemos para estimar la densidad predictiva de dichas predicciones empleando para ello diferentes distribuciones paramétricas, como son la Normal

Truncada, la Normal Censurada, la Beta y la distribución de Máxima Entropía (Jaynes, 1957). Las diferencias con trabajos anteriores es el uso de la distribución de Máxima Entropía que nos permitirá emplear momentos de hasta orden 4 y que la estimación de dichos momentos es realizada de forma recursiva y con parámetros que cambian con el tiempo.

4. Estimación de los momentos condicionados

Para realizar una estimación paramétrica de la densidad predictiva será necesario realizar la estimación de los momentos condicionados. La estimación de los momentos condicionados será efectuada mediante modelos de regresión en los cuales los regresores serán las predicciones puntuales de la potencia eólica.

Entonces, el modelo de regresión empleado para estimar los momentos condicionados es

$$e_{t+h|t}^M = \alpha_{0,t+h} + \alpha_{1,t+h}\hat{p}_{t+h|t} + \alpha_{2,t+h}\hat{p}_{t+h|t}^2 + \alpha_{3,t+h}\hat{p}_{t+h|t}^3 + v_{t+h}, \quad (1)$$

donde $M = 1, 2, \dots$. De esta manera para diferentes valores de M se ajustan modelos para diferentes momentos, puesto que el segundo momento está asociado al error de predicción al cuadrado y así sucesivamente.

Suponiendo que las predicciones de energía eólica han sido estimadas empleando un modelo correcto, se espera que la media condicionada del error de predicción sea cero, es decir, $m_{t+h|t}^{(0)} = E[e_{t+h|t}] = 0$. En este caso, los momentos centrales condicionados serán

$$m_{t+h}^{(M)} = E_t \left[e_{t+h|t}^M \right], \quad (2)$$

es decir, son calculados directamente empleando el modelo de regresión descrito en la expresión (1) para diferentes valores de M . No obstante, en la práctica será fácil que un predictor de la potencia eólica sea condicionalmente sesgado durante ciertos periodos de tiempo, es decir, $m_{t+h|t}^{(0)} \neq 0$. Por ello, se realiza una corrección del sesgo, por lo que para calcular los momentos centrales se emplea el modelo

$$\tilde{e}_{t+h|t}^M = \alpha_{0,t+h} + \alpha_{1,t+h}\hat{p}_{t+h|t} + \alpha_{2,t+h}\hat{p}_{t+h|t}^2 + \alpha_{3,t+h}\hat{p}_{t+h|t}^3 + v_{t+h}, \quad (3)$$

donde $\tilde{e}_{t+h|t}^M = \left(e_{t+h|t} - m_{t+h|t}^{(0)} \right)^M$ para $M = 2, 3, \dots$

Para comprobar la relación existente entre los momentos centrales de la potencia eólica generada por un parque respecto a sus predicciones se calcula la media de la estimación de los momentos para diferentes horizontes de predicción $h = 1, \dots, 48$. En la figura 3 se ve que los intervalos de confianza Bootstrap de la estimación nos confirman que los momentos siguen un patrón claro en función del valor previsto, por lo que, aparentemente, las predicciones de potencia serán buenos regresores.

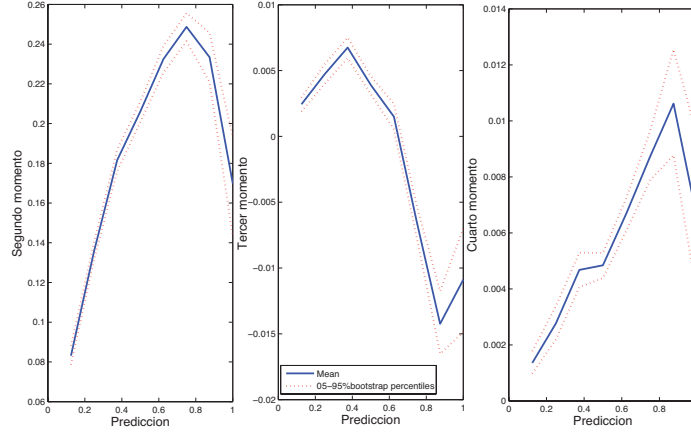


Figura 3: Media de la estimación de los momentos para diferentes horizontes

4.1. Método de estimación recursiva

El objetivo del método propuesto es ir calculando la senda de distribuciones predictivas f_{t+h} para un horizonte h , es decir, habrá que ir estimando de manera recursiva los momentos condicionados. Para ello se ha desarrollado un procedimiento adaptativo.

La estimación del modelo (3) es realizada de manera adaptativa empleando el método de mínimos cuadrados recursivos. Por simplificar la notación reescribimos (3) en forma matricial, tal que

$$y_{t+h} = X'_{t+h} \alpha_{t+h} + v_{t+h}, \quad (4)$$

donde el vector $y_{t+h} = \tilde{e}_{t+h|t}^M$, la matriz $X_{t+h} = (1, \hat{p}_{t+h|t}, \hat{p}_{t+h|t}^2, \hat{p}_{t+h|t}^3)'$ contiene las variables explicativas y el vector $\alpha_{t+h} = (\alpha_{0,t+h}, \alpha_{1,t+h}, \alpha_{2,t+h}, \alpha_{3,t+h})'$ contiene los parámetros a estimar. Entonces, empleando el método de mínimos cuadrados ponderados de manera recursiva la estimación de la expresión (4) será

$$\hat{\alpha}_{t+h} = \hat{\alpha}_{t+h-1} + M_{t+h}^{-1} X_{t+h} \hat{v}_{t+h}, \quad (5)$$

donde $\hat{v}_{t+h} = y_{t+h|t} - X'_{t+h} \hat{\alpha}_{t+h-1}$ es el error de predicción un paso adelante y la matriz de ganancia M_{t+h}^{-1} se calcula de manera recursiva según

$$M_{t+h}^{-1} = \frac{1}{\lambda_{t+h}} \left(M_{t+h-1}^{-1} - \frac{M_{t+h-1}^{-1} X_{t+h} X'_{t+h} M_{t+h-1}^{-1}}{\lambda_{t+h} + X'_{t+h} M_{t+h-1}^{-1} X_{t+h}} \right), \quad (6)$$

donde λ_{t+h} es el factor de olvido que permitirá que la estimación recursiva vaya cambiando a lo largo del tiempo. El factor de olvido empleado en este trabajo es el propuesto por Sánchez (2006), que está basado en la distancia de Cook.

Como la variable aleatoria con la que trabajamos es la potencia eólica hemos de acotar su estimación para no obtener valores fuera del intervalo $[0, 1]$. Es fácil comprobar que las acotaciones será

$$\hat{e}_{t+h|t}^M \in \begin{cases} \left[-\hat{p}_{t+h|t}^M, (1 - \hat{p}_{t+h|t})^M \right] & , \text{ si } M \text{ es impar} \\ \left[0, (1 - \hat{p}_{t+h|t})^M \right] & , \text{ si } M \text{ es par} \end{cases}. \quad (7)$$

En la figura 4 podemos ver la senda de estimaciones de los 4 primeros momentos centrales durante un periodo de tiempo, para diferentes horizontes de predicción. Además en la figura 5 podemos ver la asimetría y curtosis asociadas a dichos momentos centrales. En alguno de los gráficos hemos tenido que emplear diferentes escalas en los ejes de las ordenadas, para facilitar ver la evolución de los momentos, puesto que según aumenta el horizonte de predicción, los valores serán mayores.

5. Modelización adaptativa de densidades predictivas

En la sección 2.1 mostramos que la distribución de la potencia tiene varias características que deberemos de tener en cuenta a la hora de seleccionar las distribuciones con las que trabajaremos, como son que es acotada, que es unimodal, que es condicionalmente asimétrica, etc. Como la potencia producida está acotada no se puede emplear la distribución Normal, por ello hemos empleado dos posibles opciones basadas en la Normal como son la Normal Truncada y la Normal Censurada. Estas distribuciones normales con soporte acotado presentan ya diferentes asimetrías y curtosis que podrían ser adecuadas para nuestro problema.

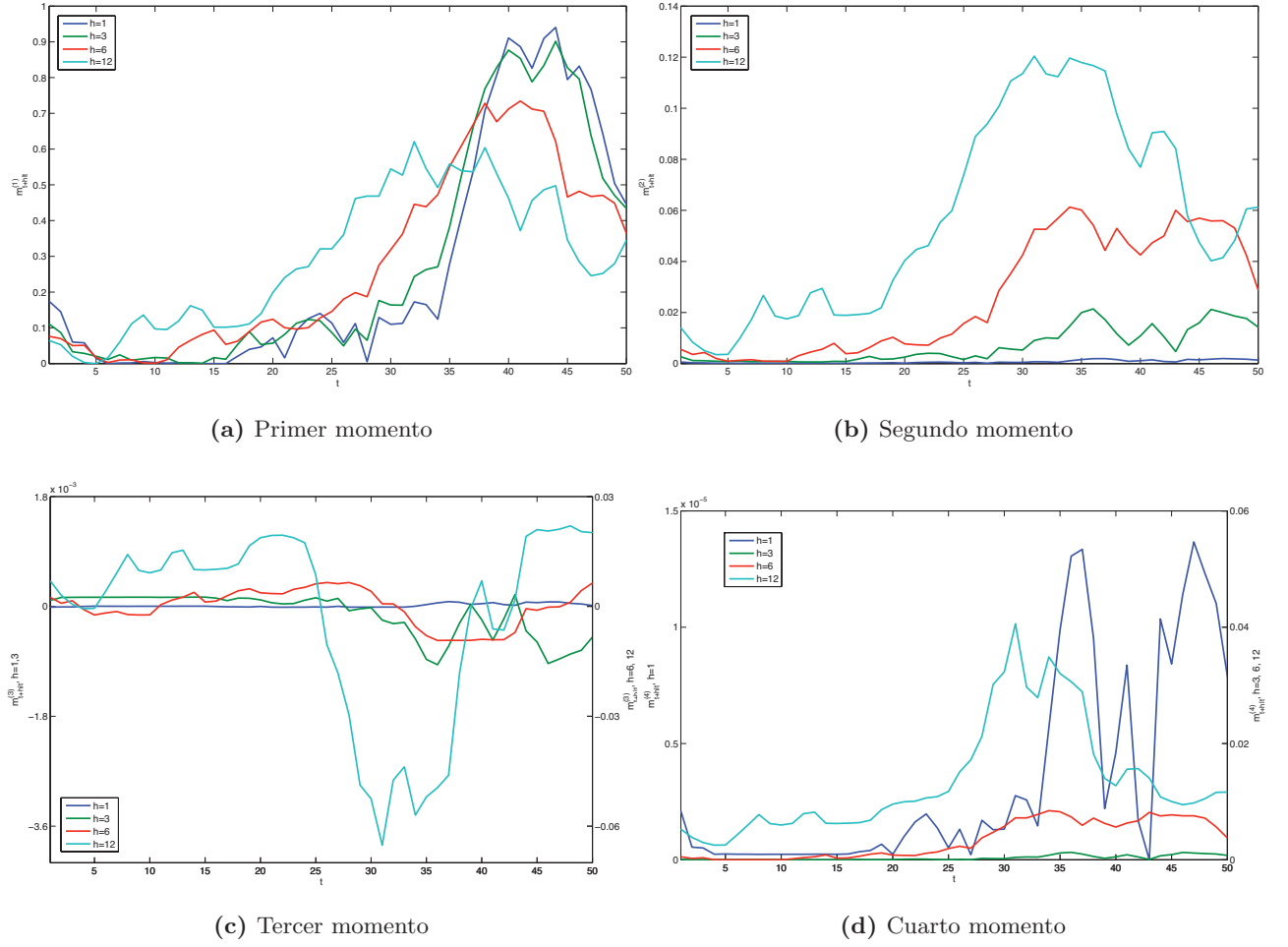


Figura 4: Estimación recursiva de los momentos centrados condicionados para $h=1,3,6,12$.

Otra distribución apropiada para nuestros datos será la Beta, ya que está acotada en $[0, 1]$ y se puede ajustar a fenómenos asimétricos. Por último, emplearemos la distribución de Máxima Entropía, ya que también nos permitirá obtener una densidad acotada y nos es muy flexible para adaptarse a momentos condicionados cambiantes.

5.1. Distribución Normal truncada

Sea $\phi(\cdot)$ la función de densidad de una distribución Normal de media μ y desviación σ y $\Phi(\cdot)$ su correspondiente función de distribución. Entonces, la variable $y \in R = [a, b]$, donde $-\infty < a < b < \infty$, sigue una distribución Normal Truncada de densidad

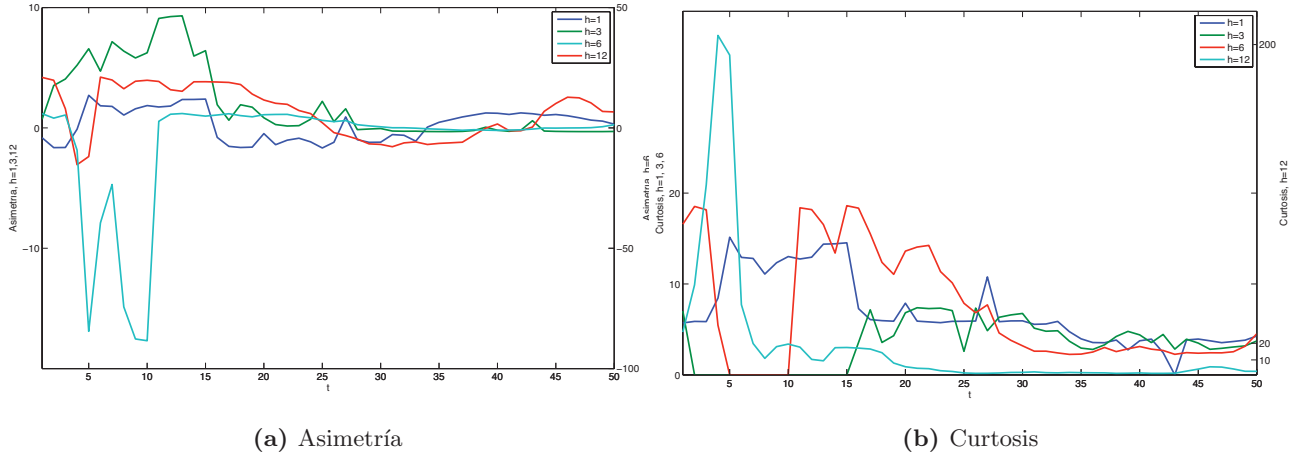


Figura 5: Estimación recursiva de los momentos condicionados para $h=1,3,6,12$.

$$f(y) = \frac{\frac{1}{\sigma} \phi\left(\frac{y-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}, \quad a < y < b. \quad (8)$$

Entonces, los momentos de la variable truncada podrán ser calculados fácilmente a partir de los momentos de la variable sin truncar μ y σ (p.ej. Cohen 1961),

$$\mu_{(NT)} = \mu + \sigma\delta; \quad (9)$$

$$\sigma_{(NT)}^2 = \sigma^2 \left[1 - \delta^2 - \frac{[a - \mu] \phi\left(\frac{a-\mu}{\sigma}\right) - [b - \mu] \phi\left(\frac{b-\mu}{\sigma}\right)}{\sigma \left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \right]} \right], \quad (10)$$

donde

$$\delta = \frac{\phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}. \quad (11)$$

5.1.1. Estimación de la densidad predictiva

Empleando la expresión (8) se puede estimar la función de densidad Normal Truncada empleando para ello los momentos de la variable sin truncar. El problema es que calcular los momentos de la variable truncada a partir de los de la variable sin truncar, como vimos en (9) y (10), es directo, pero la dirección contraria, que justamente es la que necesitamos, no lo es.

Si asumimos que la distribución de las predicciones empleadas en nuestra estimación es una Normal Truncada, estimamos los dos primeros momentos condicionados empleando el modelo (3), con $M = 1, 2$. El problema es que de esta manera obtenemos los momentos de la variable truncada $\mu_{(NT)}$ y $\sigma_{(NT)}$ y no los momentos de la variable sin truncar que necesitaremos para estimar la densidad $f_{t+h}^{(NT)}$.

Cohen (1991) muestra que los momentos sin truncar pueden ser estimados mediante

$$\hat{\sigma} = \frac{b - a}{\hat{\xi}_2 - \hat{\xi}_1}; \quad (12)$$

$$\hat{\mu} = a - \hat{\sigma}\hat{\xi}_2, \quad (13)$$

donde

$$\xi_1 = \frac{a - \mu}{\sigma} \text{ y } \xi_2 = \frac{b - \mu}{\sigma}. \quad (14)$$

La estimación de $\hat{\xi}_1$ y $\hat{\xi}_2$ puede ser efectuada mediante diferentes métodos iterativos como el de Newton-Raphson, pero Cohen (1991) propone emplear un procedimiento gráfico desarrollado en Cohen (1957).

Entonces, empleando las expresiones (9) y (10) se obtienen los momentos sin truncar μ y σ resolviendo el sistema mediante un algoritmo de búsqueda. De esta manera, empleamos la expresión (8) para obtener la estimación de la distribución predictiva $f_{t+h}^{(NT)}$. En la figura (6) podemos ver un ejemplo de la senda de estimaciones $f_{t+1}^{(NT)}$ para diferentes periodos de tiempo.

5.2. Distribución Normal censurada

Sea y una variable aleatoria con función de densidad $\phi(\cdot)$ Normal de media μ y desviación σ y $\Phi(\cdot)$ su correspondiente función de distribución. Entonces, la variable $y' \in R = [a, b]$, donde $-\infty < a < b < \infty$, sigue una distribución Normal Censurada donde

$$y' = \begin{cases} a & , \text{ si } y \leq a \\ N(\mu, \sigma) & , \text{ si } a < y < b \\ b & , \text{ si } y \geq b \end{cases} \quad (15)$$

En este caso, las colas de la Normal que sobrepasen el intervalo en el que la variable está acotada se acumulan en los extremos. Los momentos de la variable censurada son (Cohen, 1991)

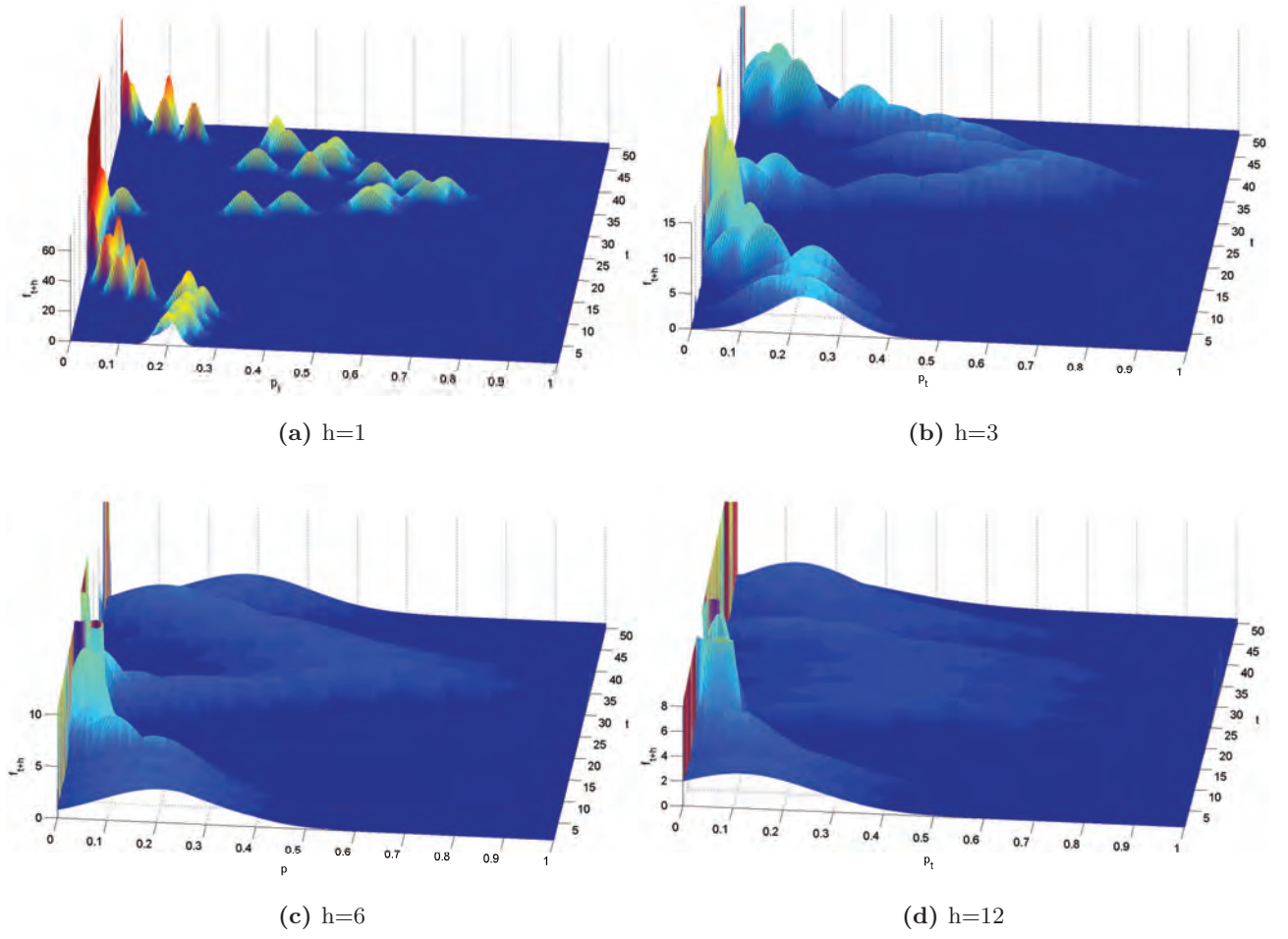


Figura 6: Ejemplo de la senda de densidades predictiva empleando la distribución Normal Truncada para diferentes horizontes

$$\mu_{(NC)} = a\Phi(\xi_1) + [\Phi(\xi_2) - \Phi(\xi_1)](\mu + \sigma\delta) + [1 - \Phi(\xi_2)]b. \quad (16)$$

$$\begin{aligned} \sigma_{(NC)}^2 &= a^2\Phi(\xi_1) + [\Phi(\xi_2) - \Phi(\xi_1)](\sigma^2 + \mu^2 + 2\mu\sigma\delta) + \dots \\ &\quad \sigma[[a - \mu]\Phi(\xi_1) - [b - \mu]\Phi(\xi_2)] + [1 - \Phi(\xi_2)]b^2 - \mu_{(NC)}^2, \end{aligned} \quad (17)$$

donde

$$\delta = \frac{\phi(\xi_2) - \phi(\xi_1)}{\Phi(\xi_2) - \Phi(\xi_1)}. \quad (18)$$

5.2.1. Estimación de la densidad predictiva

En este caso, nos encontramos en el mismo problema que para la estimación de la distribución truncada. Además, Cohen (1991) muestra como la estimación de los momentos de la variable sin censurar se hará mediante las mismas expresiones que en la distribución Normal Truncada, es decir,

$$\hat{\sigma} = \frac{b - a}{\hat{\xi}_2 - \hat{\xi}_1}; \quad (19)$$

$$\hat{\mu} = a - \hat{\sigma}\hat{\xi}_2, \quad (20)$$

solo que la estimación de $\hat{\xi}_1$ y $\hat{\xi}_2$, en esta ocasión será diferente. Para hallar los momentos de la variable sin censurar, dicha estimación tendrá que ser realizada mediante un procedimiento iterativo descrito en Cohen (1957, 1991).

El primer paso importante en un proceso iterativo es obtener una buena estimación inicial que esté próxima a la solución óptima, para que de esta manera el número de iteraciones sea el menor posible. Como valores iniciales $\hat{\xi}_1^0$ y $\hat{\xi}_2^0$, Cohen (1957) emplea

$$\Phi(\hat{\xi}_1^0) = \frac{c_1}{N} \text{ y } 1 - \Phi(\hat{\xi}_2^0) = \frac{c_2}{N}, \quad (21)$$

donde c_1 y c_2 son las el número de observaciones muestrales censuradas por la izquierda y por la derecha respectivamente, y N es el tamaño total de la muestra. Una vez hallados los valores iniciales Cohen (1957) recomienda mejorar la estimación realizando un par de iteraciones empleando las "sustituciones sucesivas de Cohen", donde

$$\begin{aligned} \hat{\xi}_1^{i+1} &= \frac{\hat{\xi}_2^{i+1}v_1/w - (\Omega_1^i - \Omega_2^i)}{1 - v_1/w} \\ \hat{\xi}_2^{i+1} &= (\Omega_1^i - \Omega_2^i) + \frac{\left\{ [(\Omega_1^i - \Omega_2^i)v_1/w + \Omega_2^i] - \sqrt{[(\Omega_1^i - \Omega_2^i)v_1/w + \Omega_2^i]^2 + 4s^2/w^2} \right\} w(v_1 - w)}{2s^2} \end{aligned} \quad (22)$$

donde $v_1 = \mu_{(NC)} - a$, $w = b - a$, $s^2 = \sigma_{(NC)}^2$ y

$$\begin{aligned} \Omega_1^i &= \frac{c_1}{n} \frac{\phi(\hat{\xi}_1^i)}{\Phi(\hat{\xi}_1^i)} \\ \Omega_2^i &= \frac{c_2}{n} \frac{\phi(\hat{\xi}_2^i)}{1 - \Phi(\hat{\xi}_2^i)} \end{aligned} \quad (23)$$

siendo n el número de observaciones muestrales sin censurar. A partir de entonces, se emplea un método de Newton hasta alcanzar el óptimo. Para ello, se usan los valores $\hat{\xi}_1^{i+1}$ y $\hat{\xi}_2^{i+1}$ obtenidos para estimar

los valores iniciales del método de Newton $\hat{\mu}^0$ y $\hat{\sigma}^0$, y a partir de ellos se buscarán las estimaciones de los momentos muestrales sin censurar.

Es importante notar que al estimar los momentos empleando los datos que poseemos, obtenemos los momentos correspondientes o bien a la muestra truncada o bien a la censurada, dependiendo del caso en el que estemos. Por ello, es necesario resolver el sistema que nos permita calcular el valor de los momentos μ y σ de la distribución sin truncar o sin censurar. No es infrecuente encontrar trabajos donde se ignora este problema.

En la figura 7 podemos ver un ejemplo de la senda de estimaciones $f_{t+1}^{(NC)}$ para diferentes periodos de tiempo.

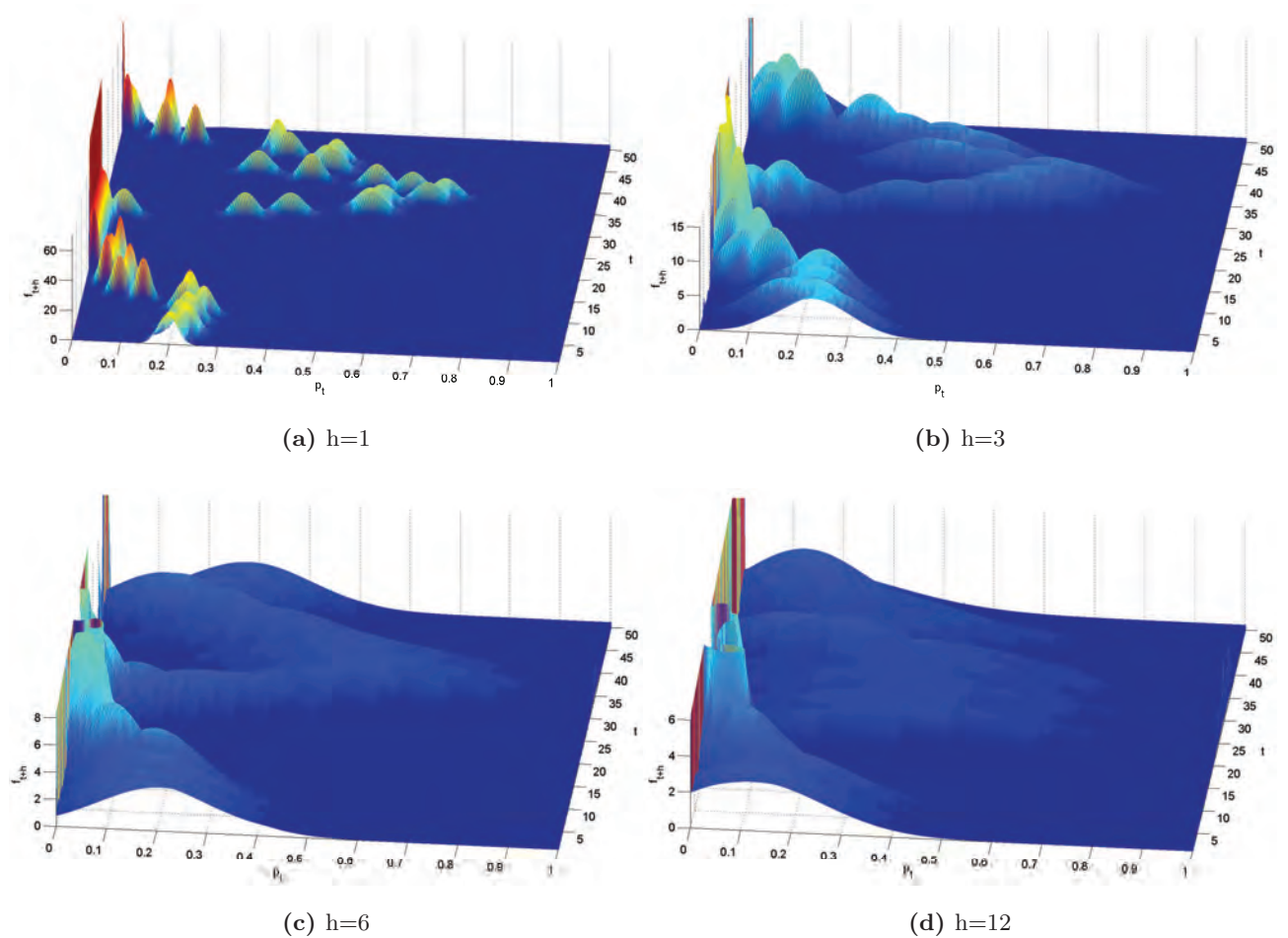


Figura 7: Ejemplo de la senda de densidades predictiva empleando la distribución Normal Censurada para diferentes horizontes

5.3. Distribución Beta

Sea y una variable que sigue una distribución $Beta(\alpha, \beta)$, entonces su función de densidad es

$$f(y) = \frac{y^{\alpha-1} [1-y]^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < y < 1, \quad (24)$$

donde $B(\alpha, \beta)$ es la función Beta. Los parámetros α y β se estimarán, por ejemplo, empleando los dos primeros momentos de la variable y a partir de la relación

$$\alpha = \frac{\mu^2 - \mu\sigma^2}{\sigma^2 - \mu^2}; \quad \beta = \frac{\mu - \sigma^2}{\sigma^2 - \mu^2} - \alpha \quad (25)$$

5.3.1. Estimación de la densidad predictiva

En este caso la estimación será sencilla, ya que empleando la expresión (3), con $M = 1, 2$, obtendremos la estimación de los dos primeros momentos, que nos permiten calcular los parámetros de la distribución Beta mediante la expresión (25). Entonces, la función de densidad Beta $f_{t+h}^{(B)}$ se calcula empleando (24). En la figura 8 podemos ver un ejemplo de la senda de estimaciones $f_{t+1}^{(B)}$ para diferentes periodos de tiempo.

5.4. Distribución de Máxima Entropía

La entropía puede ser vista como una medida de la incertidumbre para una función de densidad (Shannon, 1948)

$$S = - \int_R f(y) \ln(f(y)) dy. \quad (26)$$

Jaynes (1957) empleó este concepto para proponer una distribución de probabilidad que minimice la información previa necesaria, que será aquella que maximice la entropía de Shannon. Entonces, la función de densidad de máxima entropía $f(y)$ definida en el intervalo $[a, b]$ es aquella que cumpla

$$\begin{aligned} \text{máx} \quad & - \int_a^b f(y) \ln(f(y)) dy \\ \text{s.a.} \quad & \int_a^b f(y) dy = 1 \\ & \int_a^b y^i f(y) dy = m_i, \quad i = 1, \dots, M, \end{aligned} \quad (27)$$

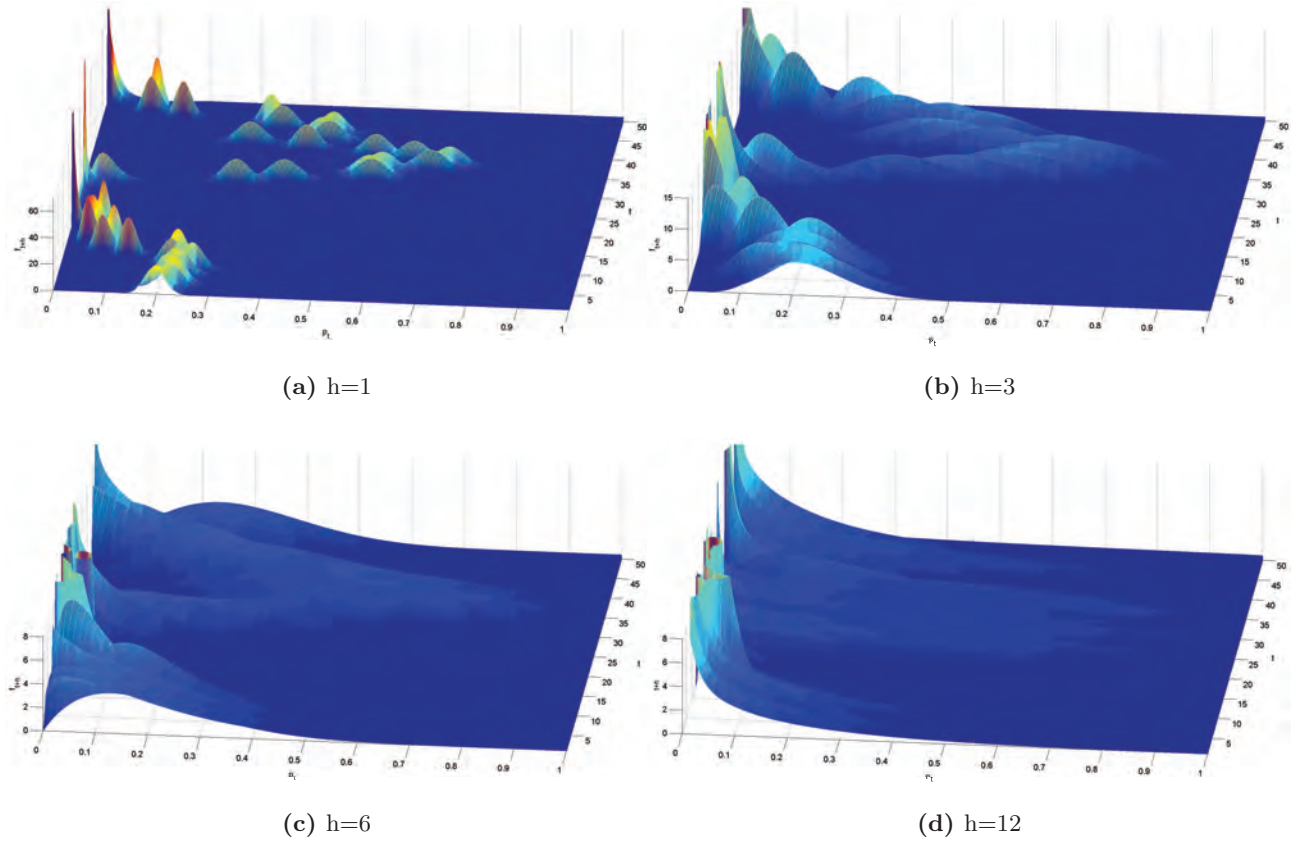


Figura 8: Ejemplo de la senda de densidades predictiva empleando la distribución Beta para diferentes horizontes

donde los m_i son los M primeros momentos respecto al origen. Jaynes (1957) muestra que la forma analítica de la función de densidad de máxima entropía $f(y)$ es

$$f(y) = \exp \left[\lambda_0 + \sum_{i=1}^M \lambda_i y^i \right], \quad (28)$$

donde

$$\lambda_0 = -\ln \left[\int_a^b \exp \left(\sum_{i=1}^M \lambda_i y^i \right) dy \right], \quad (29)$$

y los multiplicadores $\lambda_1, \dots, \lambda_M$ son calculados resolviendo el sistema de M ecuaciones no lineales

$$\left[\int_a^b y^j \exp \left(\sum_{i=1}^M \lambda_i y^i \right) dy \right] \left[\int_a^b \exp \left(\sum_{i=1}^M \lambda_i y^i \right) dy \right]^{-1} = m_j, \quad j = 1, \dots, M. \quad (30)$$

El sistema descrito en (30) solo tiene solución analítica cuando $M = 1$. En este trabajo llegaremos a emplear un $M = 4$, por lo que el sistema (30) tendrá que ser resuelto numéricamente. El procedimiento que hemos seguido para ello es el descrito por Siddall y Diab (1975). Los pasos del procedimiento son:

1. Calcular los M primeros momentos condicionados centrales de manera adaptativa.
2. Transformar los M momentos centrales m_j en momentos en el origen $m_j^{(0)}$, empleando

$$\begin{aligned} m_1^{(0)} &= m_1 \\ m_j^{(0)} &= m_j - m_1^j (-1)^j - \sum_{k=1}^{j-1} \frac{(-1)^k j!}{k!(j-k)! m_1^k m_{j-k}^{(0)}} \end{aligned} \quad (31)$$

3. Elegir los valores iniciales $\lambda_1^0, \dots, \lambda_M^0$. Siddall y Diab (1975) proponen diferentes alternativas, entre las que hemos escogido la más conservadora $\lambda_1^0 = \dots = \lambda_M^0 = 0$.
4. Calcular los multiplicadores $\lambda_1, \dots, \lambda_M$ resolviendo el problema de programación no lineal sin restricciones

$$1 - \frac{\sum_{i=1}^l a_i x_i^j \exp\left(\sum_{j=1}^M \lambda_j x_i^j\right)}{m_j^0 \sum_{i=1}^l a_i \exp\left(\sum_{j=1}^M \lambda_j x_i^j\right)} = R_j, \quad j = 1, \dots, M, \quad (32)$$

donde a_i son los multiplicadores de integración numérica, l es el número de intervalos empleados en dicha integración numérica y R_j son los residuos normalizados empleados en la resoluciones de las ecuaciones simultáneas (32).

5. Emplear la expresión (29) para calcular λ_0 .
6. Emplear la expresión (28) para estimar la distribución de máxima entropía.

5.4.1. Estimación de la densidad predictiva

El método de máxima entropía ha sido utilizado para la estimación de la densidad de la velocidad del viento en trabajos previos (Ramírez y Carta, 2006; Akpinar y Kavak Akpinar, 2007; Liu y Chang, 2010). Sin embargo no ha sido empleado para la modelización de la densidad predictiva de la potencia eólica, a pesar de que su uso tiene varias ventajas para dicho problema. La principal, es que basándonos sólo en los datos y sin necesidad de ningún tipo de asunción previa, podemos estimar una distribución acotada en el dominio $[a, b]$ que queramos.

Para estimar la densidad predictiva $f_{t+h}^{(MEP)}$ estimaremos los cuatro primeros momentos condicionados empleando la expresión (3), con $M = 1, 2, 3, 4$. Entonces, se resuelve de manera numérica el sistema descrito en (30), y obtenemos la estimación de la densidad predictiva $f_{t+h}^{(MEP)}$ mediante (28) y (29). En la figura 9 podemos ver un ejemplo de la senda de estimaciones $f_{t+1}^{(MEP)}$ para diferentes periodos de tiempo.

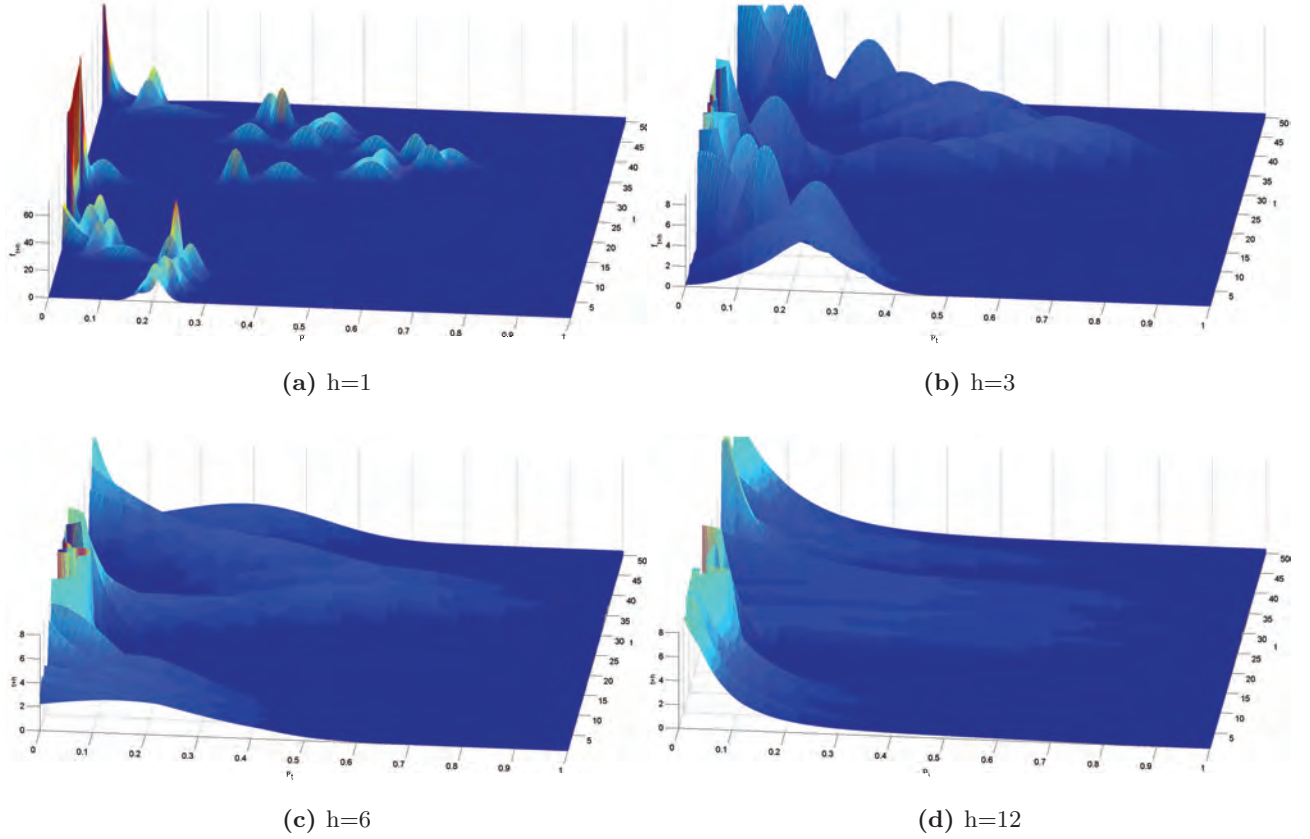


Figura 9: Ejemplo de la senda de densidades predictiva empleando la distribución de máxima entropía para diferentes horizontes

5.5. Comparación de las distribuciones empleadas

En la figura 10 podemos ver varios ejemplos de las diferentes estimaciones obtenidas en un determinado instante de tiempo para diferentes horizontes de predicción. Para $h = 1$ vemos un ejemplo en el que todas las distribuciones funcionan de manera similar puesto que nos encontramos en mitad del intervalo y la desviación es pequeña. En el ejemplo para un $h = 3$ vemos las diferencias existen-

tes cuando nos encontramos en un extremo. Diferencias que se irán acrecentando según aumenta el horizonte y con ello la desviación como vemos en los gráficos restantes.

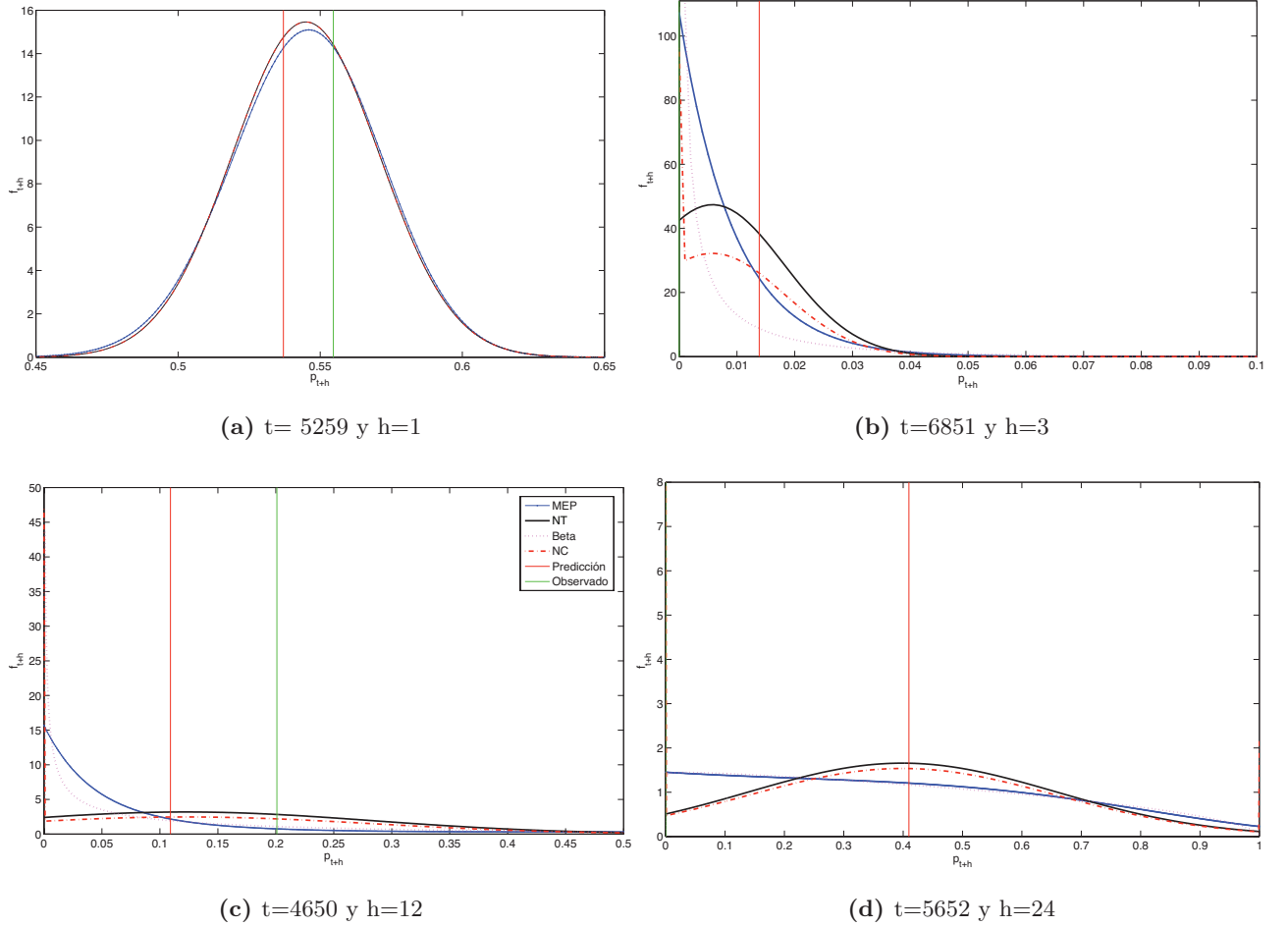


Figura 10: Densidades predictivas para diferentes instantes de tiempo y horizontes

6. Criterios de evaluación de densidades predictivas

Existen multitud de herramientas para evaluar el correcto funcionamiento la una secuencia de estimaciones de densidades predictivas. Se pueden clasificar los diferentes métodos en dos tipos distintos, por una parte están aquellos métodos que sirven para comprobar que la densidad predictiva estimada es correcta. Estos son métodos necesarios, pero no suficientes pues no garantizan un buen comportamiento predictivo futuro. Por ello son necesarios otro tipo de métodos que comparen entre diferentes densidades predictivas y digan cuál es la más eficiente.

6.1. Comprobación de la estimación

La evaluación de una densidad predictiva se basa en el "*Prequential principle*", llamado así por Dawid (1984). Dawid formó la palabra *Prequential* uniendo los términos secuencia probabilística (*Probabilistic sequential*). La idea del principio es que dado un proceso $\{Y_{t+h}\}$, y una secuencia de densidades predictivas $\{f_{t+h}\}$, la eficiencia de la estimación de las densidades predictivas podrá ser evaluada empleando únicamente el par $\{Y_{t+h}, f_{t+h}\}$.

Empleando este principio Dawid (1984) propone evaluar si una secuencia de densidades predictivas es adecuada empleando el método PIT (*Probability Integral Transform*). Sean u_t los valores PIT, definidos como

$$u_{t+h} = \int_{-\infty}^{Y_{t+h}} f_{t+h} dY_t, \quad (33)$$

si asumimos que la secuencia de estimaciones $\{f_{t+h}\}$ es adecuada entonces, Rosenblatt (1952) probó que los valores PIT $\{u_{t+h}\}$ serán un proceso independiente e idénticamente distribuido según una distribución Uniforme $U[0, 1]$. El método PIT propuesto por Dawid (1984) empezó a ser ampliamente usado a partir del trabajo de Diebold et al. (1998) que popularizó su uso. A partir de este trabajo diferentes propuestas para evaluar diferentes propiedades estadísticas de la secuencia de densidades predictivas se han ido proponiendo. En Corradi y Swanson (2006) se hace una amplia revisión. Más recientemente González-Rivera et al. (2011) proponen una ampliación para el caso multivariante.

Durante este trabajo como herramienta para evaluar la secuencia de estimaciones predictivas hemos empleado el método PIT. Para comprobar que los valores PIT siguen una distribución uniforme hemos realizado, como se hace habitualmente, el histograma de los valores PIT obtenidos. Otra posible representación, quizás más útil para comparar secuencias de densidades predictivas diferentes es obtener un gráfico de los percentiles empíricos del proceso $\{u_{t+h}\}$ obtenido respecto a los percentiles nominales esperados.

Hamill (2001) muestra un ejemplo del funcionamiento del método PIT como método de evaluación. En el ejemplo, emplea cuatro densidades predictivas diferentes. Una es la ideal y las otras tres son densidades sesgadas. En todos los casos, el método PIT nos dice que las cuatro distribuciones son correctas a pesar de que tres de ellas no lo son. Esto es debido a que el método PIT es una condición necesaria pero no suficiente, por lo que se necesitan emplear otros métodos que permita distinguir cual

es el predictor ideal.

6.2. Verificación de las predicciones

El concepto de verificación de las predicciones fue introducido en el campo de la meteorología por Jolliffe y Stephenson (2003) y se basa en la idea de obtener medidas que permiten la evaluación de las densidades predictivas. La mayoría de las medidas existentes se basan en el Brier Score, introducido por Brier (1950). Sea F_t la función de distribución acumulada predictiva de la potencia eólica en el instante de tiempo t y sea p_t el valor observado en dicho instante de tiempo, entonces el Brier Score se calcula mediante

$$BS(x) = \frac{1}{N} \sum_{t=1}^N [F_t - H(x - p_t)]^2, \quad (34)$$

donde $H(x - p_t)$ es la función Heaviside y toma valores 0 cuando $x < p_t$ y valores 1 en caso contrario, siendo x el dominio de p_t . El Brier Score es uno de los métodos conocidos como *Scoring Rules*. Estos métodos asignan una puntuación numérica a la densidad predictiva, basandose en el concepto de concentración. Esto es, una distribución predictiva más concentrada, nos dará predicciones mejores. Así pues, un menor valor del Brier Score nos dará una densidad predictiva mejor.

Diferenciar cuál de las densidades predictivas alternativas es mejor puede ser complicado a partir sólo del Brier Score. Por esta razón, es aconsejable una medida que nos diga cual de las distribuciones acumulará menor valor Brier Score. Existen multitud de medidas que permiten esta tarea. En Gneiting y Raftery (2007) podemos ver una amplia revisión. En este trabajo hemos empleado la medida propuesta por Matheson y Winkler (1976), el CRPS (*Continuous Ranked Probability Score*) que vendrá definido por

$$CRPS = \int_{-\infty}^{\infty} BS(x) \, dx, \quad (35)$$

la cual, al ser una medida asociada al Brier Score, nos dará como mejor densidad predictiva a aquella que tenga un menor valor CRPS.

7. Selección de la distribución óptima empleando criterios de evaluación

En la sección anterior, vimos los resultados de las diferentes funciones de distribución que hemos considerado. Sin embargo, es posible que a lo largo del tiempo la que consideramos mejor densidad predictiva vaya cambiando entre las diferentes opciones que consideramos. Por ello, proponemos emplear los criterios de evaluación explicados anteriormente para ir escogiendo la distribución más adecuada dependiendo del instante de tiempo.

Para seleccionar en un determinado instante de tiempo la densidad predictiva más adecuada hemos tenido en cuenta dos posibles criterios:

1. Mínimo Brier Score ($f_{t+h}^{(\min BS)}$). En el instante de tiempo $t + h$ seleccionaremos como densidad predictiva aquella que tenga menor valor del Brier Score predicho en $\hat{p}_{t+h|t}$.
2. Mínimo CRPS ($f_{t+h}^{(\min CRPS)}$). En el instante de tiempo $t + h$ seleccionaremos como densidad predictiva aquella que minimice el valor de CRPS.

El problema es que para calcular tanto el Brier Score como el CRPS necesitamos el valor observado en el instante de tiempo deseado. Entonces, suponiendo que en el instante de tiempo t hemos calculado una predicción del BS y del CRPS, y dado que en dicho instante conocemos tanto ambos valores, generaremos la siguiente predicción empleando el alisado simple tipo EWMA (Holt, 1957). Es decir,

$$\widehat{BS}_{t+1} = \lambda \widehat{BS}_t + (1 - \lambda) BS_t,$$

donde $0 < \lambda < 1$ determinará el peso que le damos a la predicción anterior. Un valor muy próximo a 1 nos irá dando predicciones similares. Para el CRPS realizaremos la misma operación.

8. Comportamiento de la estimación en diferentes parques eólicos

Una vez estimadas las diferentes sendas de densidades predictivas empleando las diferentes distribuciones y los criterios de selección anteriormente detallados, pasamos a evaluar las diferentes es-

timaciones realizadas. Para ello comenzaremos viendo los resultados para uno de los parques eólicos disponibles, el Parque 1.

En las figuras 11, 12, 13 y 14 mostramos los histogramas de los valores PIT obtenidos para las diferentes distribuciones ajustadas para el Parque 1 a horizontes de predicción $h = 1, 6, 12, 24$. Como dijimos anteriormente si la distribución estimada es correcta esperamos que los valores PIT sigan una distribución uniforme, por lo que el histograma deberá de tener forma de rectángulo. Como vemos en los diferentes gráficos la distribución con mayores problemas es la Normal Truncada, puesto que acumula un gran número de valores próximos al cero. Las demás distribuciones en general, funcionan de manera correcta, especialmente la distribución MEP y las densidades estimadas mediante los criterios de selección detallados en la sección anterior.

Mediante la representación de los valores PIT podemos ver si las distribuciones estimadas son adecuadas, pero no podemos comprobar su eficacia. Para ello empleamos los métodos para la verificación de densidades predictivas. Una opción es comparar los diferentes valores que toma el Brier Score para las distintas distribuciones empleadas, siendo la mejor distribución aquella que menor Brier Score tenga. En la figura 15 vemos los resultados para diferentes horizontes de predicción. Como vemos a horizontes 1 y 6 la distribución MEP obtiene menores valores, pero a horizontes mayores los resultados son más confusos puesto que las diferentes distribuciones obtienen valores similares.

Para comprobar de manera más exacta el resultado observado con la representación del Brier Score, vemos los valores que toma el CRPS. En la figura 16a vemos como según aumenta el horizonte de predicción la diferencia entre las distintas distribuciones se va reduciendo. En general, podemos concluir que la mejor distribución empleada es la MEP, puesto que tiene el valor mínimo excepto a horizonte 24 en el cual la distribución *Min CRPS* obtiene un valor ligeramente menor.

Para resumir los resultados obtenidos en los 6 parques eólicos empleados vamos a agupar los resultados en un único gráfico por horizonte. Para ello calculamos los percentiles de los valores PIT estimados para cada distribución para cada uno de los parques y a continuación hallamos la media de los 6 parques. En la figura 17 vemos los resultados obtenidos. Como vemos los resultados son similares a los observados en los histogramas del PIT para el Parque 1, puesto que en general la distribución que más problemas tiene para acercarse a la diagonal es la Normal Truncada. Por último, en la figura 16

podemos ver los valores del CRPS obtenidos para los diferentes parques eólicos. Como vemos la mejor distribución es la MEP, puesto que, en general, obtiene el menor valor en los diferentes parques. Vemos además resultados prometedores empleando los criterios de selección de distribuciones, especialmente empleando el mínimo CRPS. Como vemos asumir un tipo de distribución a este tipo de datos no parece buena idea, puesto que dependiendo del parque la Normal Truncada, la Normal Censurada o la Beta puede dar buenos o malos resultados.

9. Conclusiones y líneas futuras

Calcular la densidad predictiva de las predicciones es siempre un problema complicado, y más aun si la variable a estudiar tiene las características de la potencia eólica. Como vemos en los resultados mostrados asumir que la potencia eólica sigue una distribución conocida no parece buena idea, por lo que el uso de una distribución como la de Máxima Entropía o el uso de un criterio que vaya seleccionando la distribución óptima en cada instante de tiempo parecen ideas que mejoran las propuestas previas. Además, la estimación de los momentos centrales es un problema básico para obtener una buena estimación de la densidad predictiva, y en las propuestas previas dicha estimación se limitaba al uso de la estimación muestral de los momentos.

Los tres puntos claves que proponemos en este tema son, por lo tanto, la estimación adaptativa de los momentos condicionados, el uso como densidad predictiva de la distribución de Máxima Entropía y el uso de criterios de selección para emplear la distribución más adecuada en un instante de tiempo determinado.

En cuanto a la estimación adaptativa, se han realizado ya diferentes pruebas con modelos semejantes empleando diferentes potencias y los resultados son similares. Una posible ampliación sería añadir como regresores a los modelos predicciones pasadas y no solo las predicciones presentes como empleamos actualmente.

La estimación de la distribución de Máxima Entropía conlleva resolver un problema de programación no lineal en cada instante de tiempo, por ello una gran mejora computacional sería obtener una forma recursiva de dicho algoritmo, puesto que al ser una distribución condicionada esta irá cambiando dependiendo del valor anterior. Otro resultado interesante, sería la obtención de un criterio

formal para que el número de momentos empleados en cada instante de tiempo fuera variando según las características de los datos.

Por último, la selección de distribuciones es, como parecen indicar los resultados, un método prometedor. El uso del mínimo CRPS como criterio de selección de distribuciones funciona de manera eficiente sin haber empleado un método de predicción especialmente refinado. Otra idea a explorar será la de realizar una combinación de diferentes distribuciones empleando el Brier Score, que nos dice que distribución es mejor en diferentes tramos del dominio de la variable.

Referencias

- [1] Akpınar, S. y Kavak Akpınar, E. (2007). Wind Energy Analysis based on Maximum Entropy Principle (MEP)-type distribution function. *Energy Conversion and Management*, **48**: 1140-1149.
- [2] Al-Awami, A.T. y El-Sharkawi, M.A. (2009). Statistical Characterization of Wind Power Output for a given Wind Power Forecast. In *North American Power Symposium*, NAPS 2009.
- [3] Bludszuweit, H., Domínguez-Navarro, J.A. y Llombart, A. (2008). Statistical Analysis of Wind Power Forecast Error. *IEEE Transactions on Power Systems*, **23**: 983-991.
- [4] Bremnes, J.B. (2006). A Comparison of a Few Statistical Models for Making Quantile Wind Power Forecasts. *Wind Energy*, **9**: 3-11.
- [5] Brier, G.W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, **78**: 1-3.
- [6] Carta, J.A., Ramirez, P. y Velázquez, S. (2009). A review of wind speed probability distributions used in wind energy analysis, Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews*, **13**: 933-955.
- [7] Clements, M.P. y Harvey, D.I. (2010). Combining probability forecasts. *International Journal of Forecasting*, **27**: 208-223.
- [8] Cohen, A.C. (1957). On the solution of estimating equations for truncated and censored samples from normal populations. *Biometrika*, **44**: 225-236.
- [9] Cohen, A.C. (1991). *Truncated and Censored Samples*. Marcel Dekker: New York.
- [10] Corradi, V. y Swanson (2006). Predictive density evaluation. In *Handbook of Forecasting*, ed. por G. Elliot, C.W.J. Granger y A. Timmerman, Elsevier Science, 197-284.
- [11] Costa A. et al. 2008. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, **12**: 1725-1744.

- [12] Dawid, A.P. (1984). The prequential approach. *Journal of the Royal Statistical Society, Series A*, **147**: 278-292.
- [13] Diebold, F.X. y Lopez, J.A. (1996). Forecast evaluation and combination. In G.S. Maddala & C.R. Rao (Eds.), *Handbook of statistics*, **14**: 241-268. North Holland: Amsterdam.
- [14] Diebold, F.X., Gunther, T., Tay, A.S. (1998). Evaluating density forecasts with applications to finance and management. *International Economic Review*, **39**: 863-883.
- [15] Epstein, E.S. (1969). Stochastic dynamic prediction. *Tellus*, **21**: 739-759.
- [16] Giebel, G., Brownsword, R. y Kariniotakis, G. (2011). The State-of-the-Art in Short-Term Prediction of Wind Power, Deliverable Report D1.1, Project Anemos, disponible en http://anemos.cma.fr/download/ANEMOS_D1.1_StateOfTheArt_v1.1.pdf.
- [17] Gneiting, T. y Raftery, A.E. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association*, **102**: 359-378.
- [18] González-Rivera, G., Senyuz, Z. y Yoldas, E. (2011). Autocontours: dynamic specification testing. *Journal of Business and economic statistics*, **29**: 186-200.
- [19] Hamill, T.M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Reviews*, **129**: 550-560.
- [20] Holt, C.C. (1957). *Forecasting trends and seasonals by exponentially weighted moving averages*, O.N.R. Memorandum, n°52, Carnegie Institute of Technology.
- [21] Jaynes, E.T. (1957). Information theory and statistical mechanics. *Physical Reviews*, **106**: 620-630.
- [22] Jolliffe, I.T. y Stephenson, D.B. (2003). *Forecast verification: a practitioner's guide in atmospheric science*. Wiley: Chichester.
- [23] Lange, M. (2005). On the Uncertainty of Wind Power Predictions: Analysis of the Forecast Accuracy and Statistical Distribution of Errors. *Journal of Solar Energy Engineering*, **127**: 177-184.

- [24] Lau, A. y McSharry, P. (2010). Approaches for Multi-Step Density Forecasts with Application to Aggregated Wind Power. *The Annals of Applied Statistics*, **4**: 1311-1341.
- [25] Liu, F.J. y Chang, T.P. (2010). Validity analysis of maximum entropy distribution based on different moment constraint for wind energy assessment. *Energy*, doi:10.1016/j.energy.2010.11.033.
- [26] Matheson, J.E. y Winkler, R.L. (1976). Scoring rules for continuous probability distributions. *Management science*, **22**: 1087-1096.
- [27] Mitchell, J. (2008). Density forecast revisions and forecast efficiency. *National Institute of Economic and Social Research*.
- [28] Moller, J.K., Nielsen, H.A. y Madsen, H. (2008). Time-adaptive quantile regression. *Computational Statistics & Data Analysis*, **52**: 1292-1303.
- [29] Nielsen, H.A., Madsen, H., Nielsen, T.S., Badger, J., Giebel, G., Landberg, L., Sattler, K. y Feddersen, H. (2004). *Wind Power Ensemble Forecasting*. Technical report, Technical University of Denmark, DTU Informatics, Kgs. Lyngby, Denmark, 2004.
- [30] Pinson, P. (2010). *On probabilistic Forecasting of Wind Power Time-Series*. Technical report, Technical University of Denmark, DTU Informatics, Kgs. Lyngby, Denmark, 2010.
- [31] Pinson, P. y Madsen, H. (2009). Ensemble-based probabilistic forecasting at Horns Rev. *Wind Energy*, **12**: 137-155.
- [32] Ramírez, P. y Carta, J.A. (2006). The use of Wind Probability Distributions derived from the Maximum Entropy Principle in the Analysis of Wind Energy. A case study. *Energy Conversion and Management*, **47**: 2564-2577.
- [33] Sánchez, I. (2006). Recursive estimation of dynamic models using Cook's distance, with application to wind energy forecast. *Technometrics*, **48**: 61-73.
- [34] Siddall, J.N. y Diab, Y. (1975). The use in probabilistic design of probability curves generated by maximizing the Shannon entropy function constrained by moments. *Journal of Industrial Engineering*, **97**: 843-852.

- [35] Tay, A.S. y Wallis, K.F. (2000). Density Forecasting: A Survey. *Journal of Forecasting*, **19**: 235-254.
- [36] Taylor, J.W., McSharry, P.E. y Buizza, R. (2009). Wind Power Density Forecasting Using Wind Ensemble Predictions and Time Series Models. *IEEE Transactions on Energy Conversion*, **24**: 775-782.

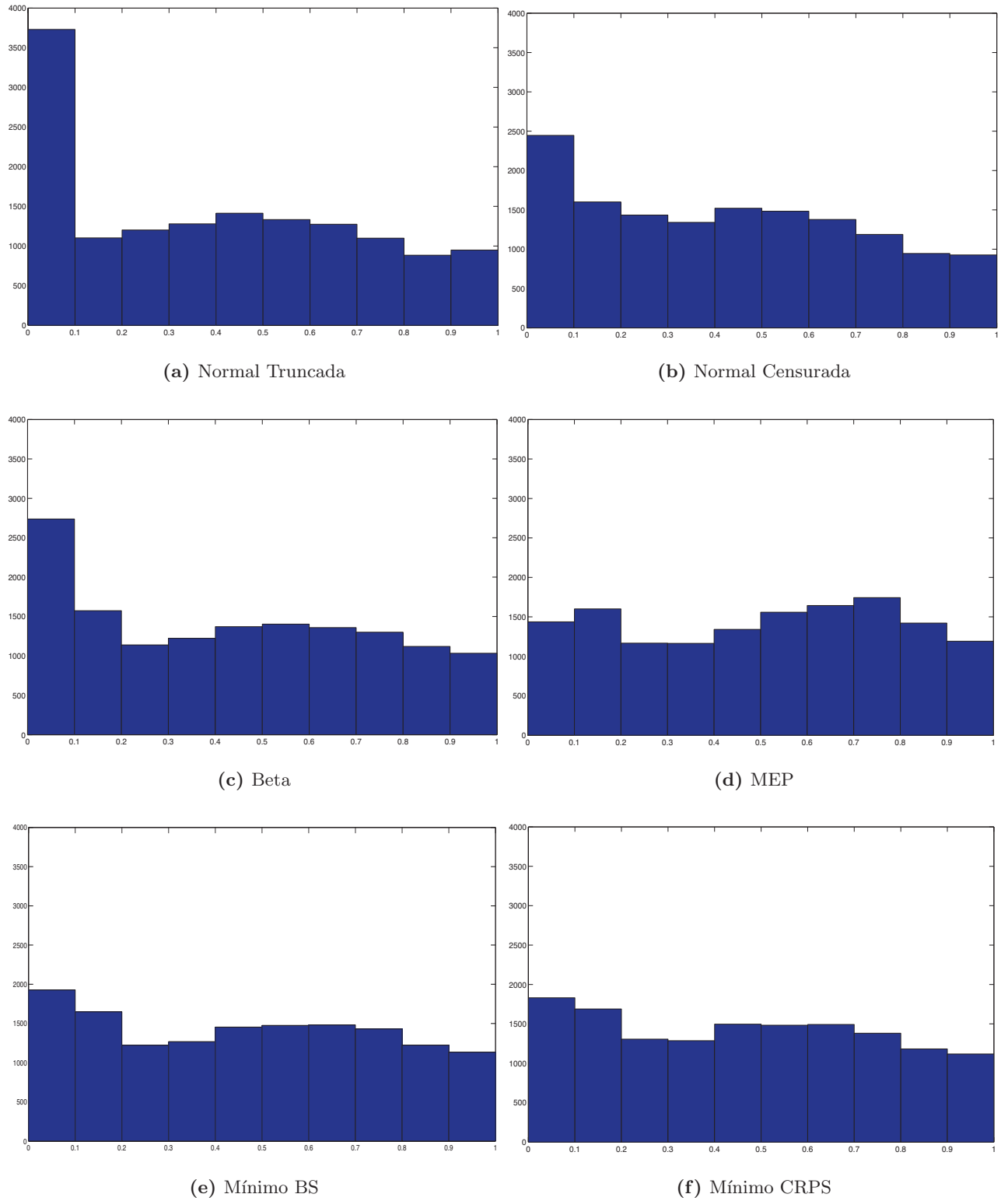
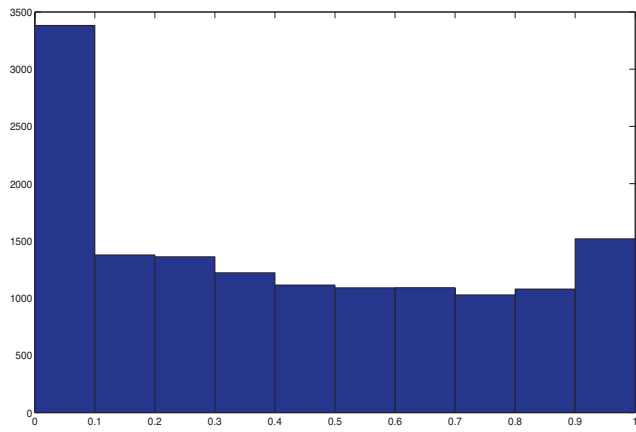
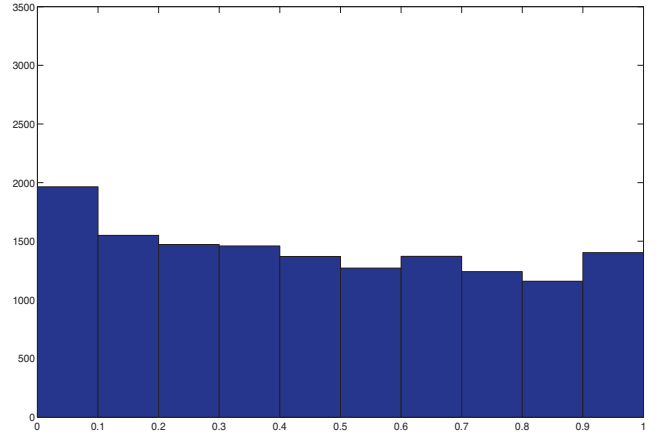


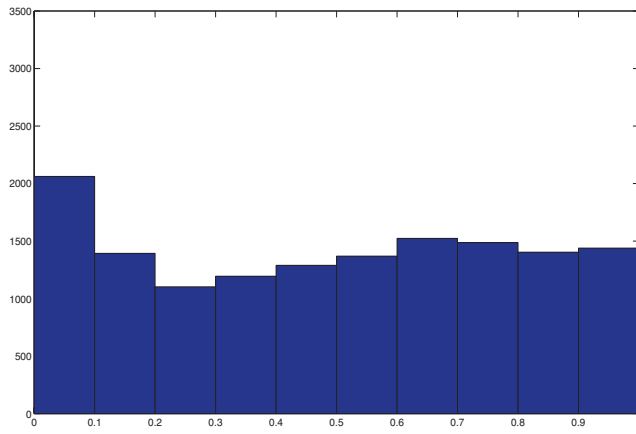
Figura 11: Histograma de los valores PIT obtenidos para diferentes densidades predictivas estimadas para un parque eólico a horizonte $h = 1$.



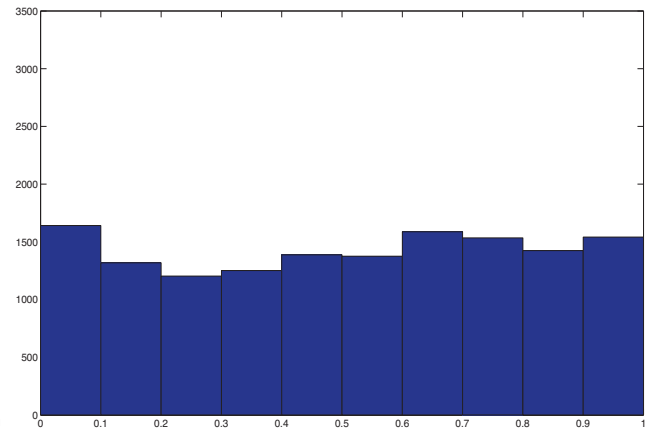
(a) Normal Truncada



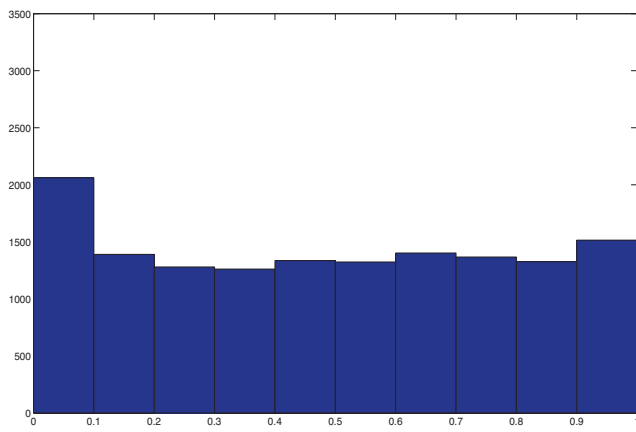
(b) Normal Censurada



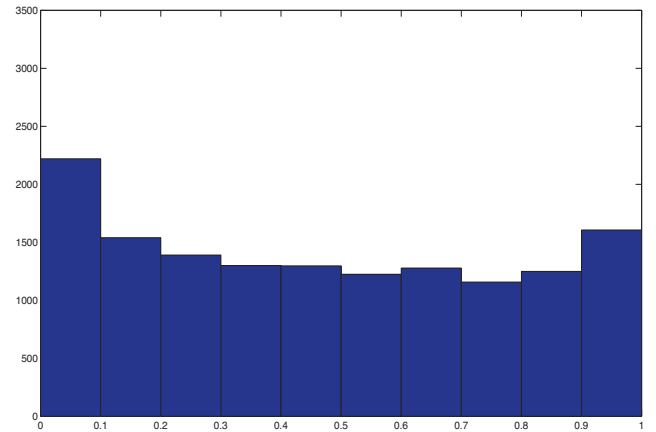
(c) Beta



(d) MEP

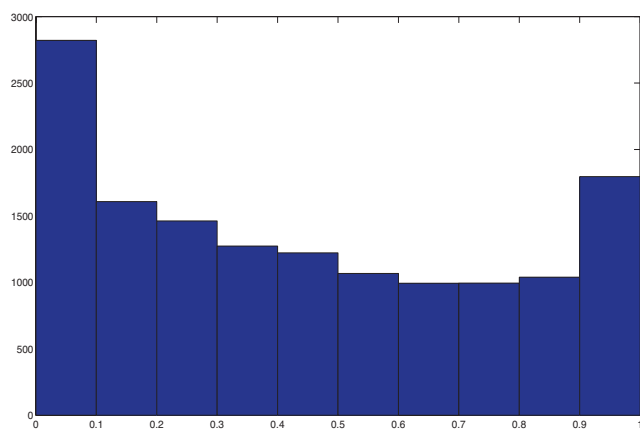


(e) Mínimo BS

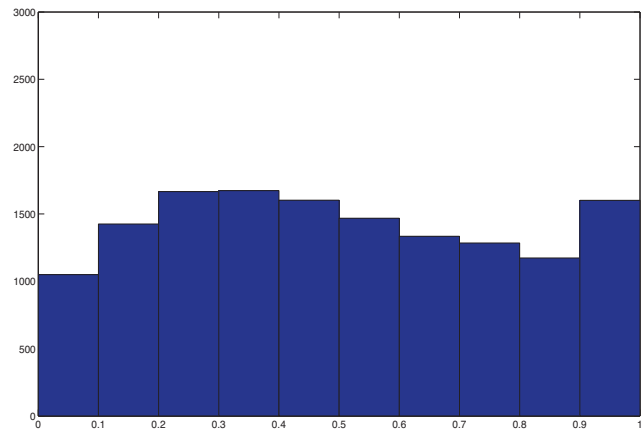


(f) Mínimo CRPS

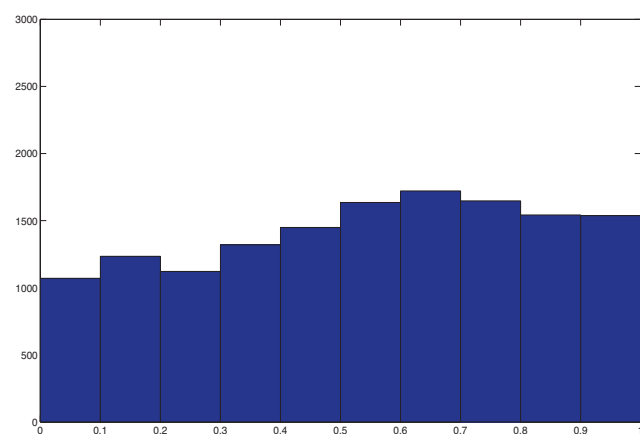
Figura 12: Histograma de los valores PIT obtenidos para diferentes densidades predictivas estimadas para un parque eólico a horizonte $h = 6$.



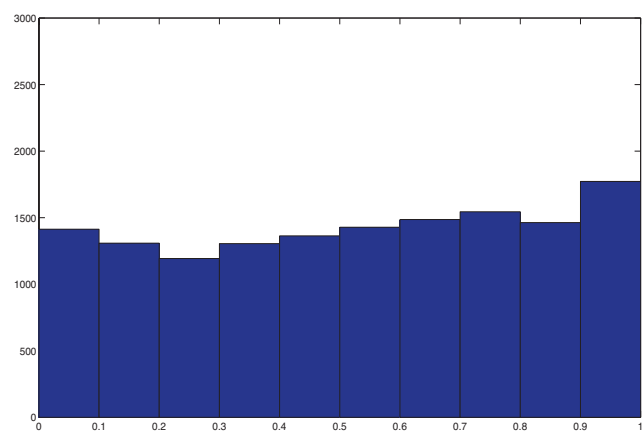
(a) Normal Truncada



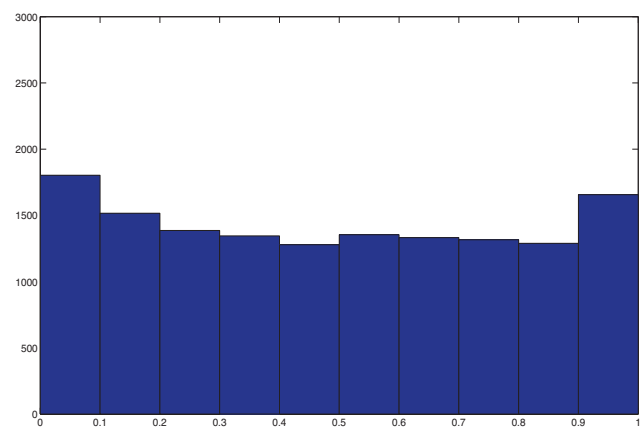
(b) Normal Censurada



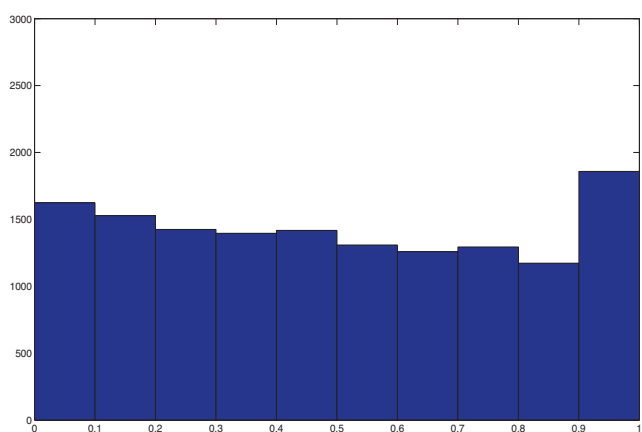
(c) Beta



(d) MEP

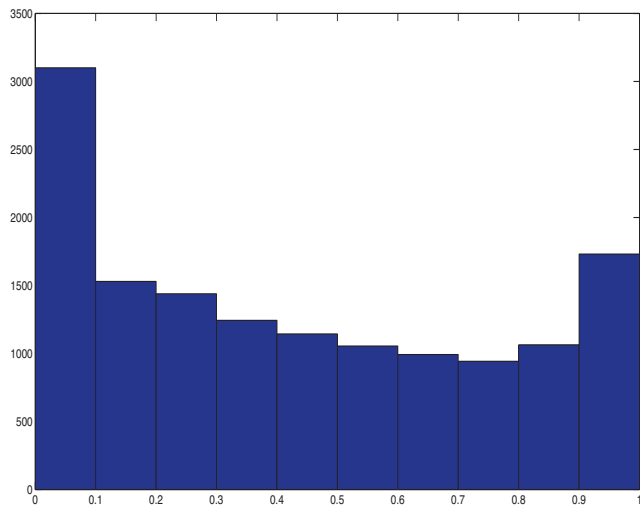


(e) Mínimo BS

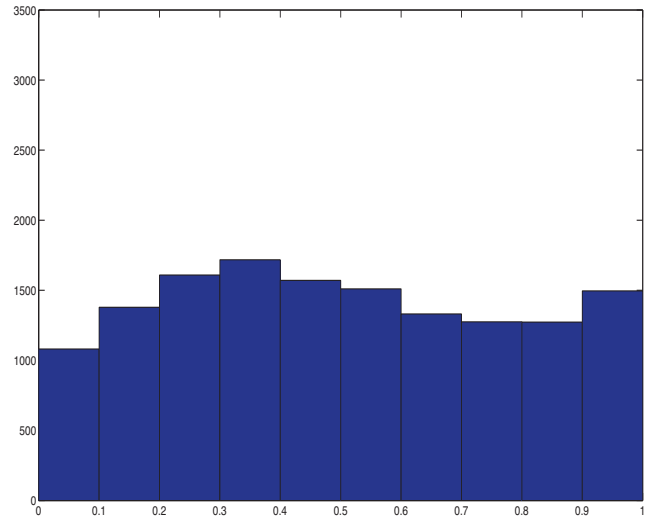


(f) Mínimo CRPS

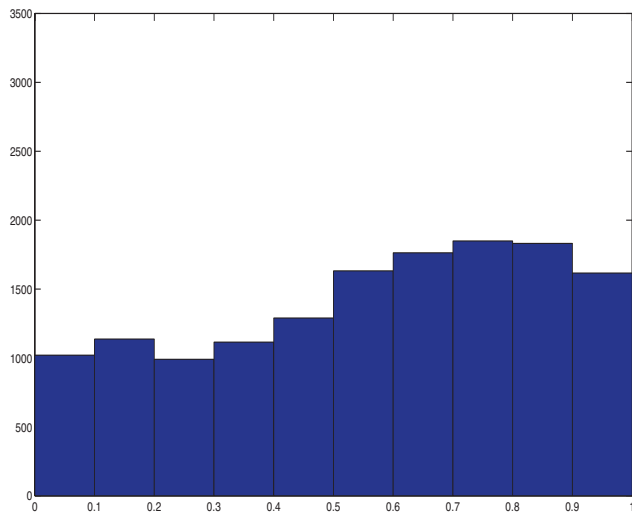
Figura 13: Histograma de los valores PIT obtenidos para diferentes densidades predictivas estimadas para un parque eólico a horizonte $h = 12$.



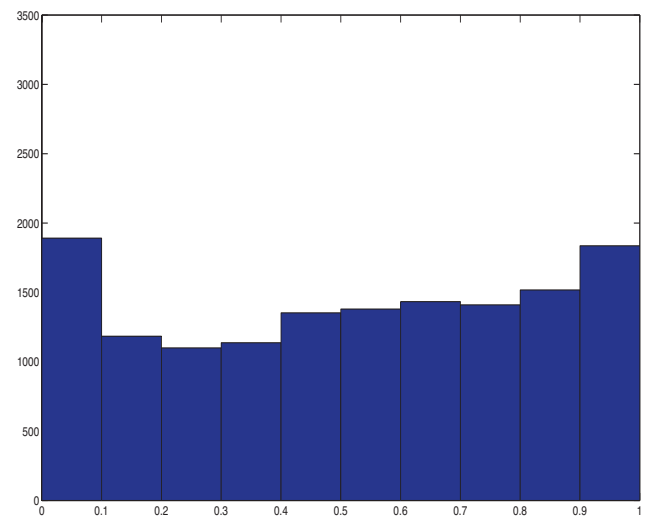
(a) Normal Truncada



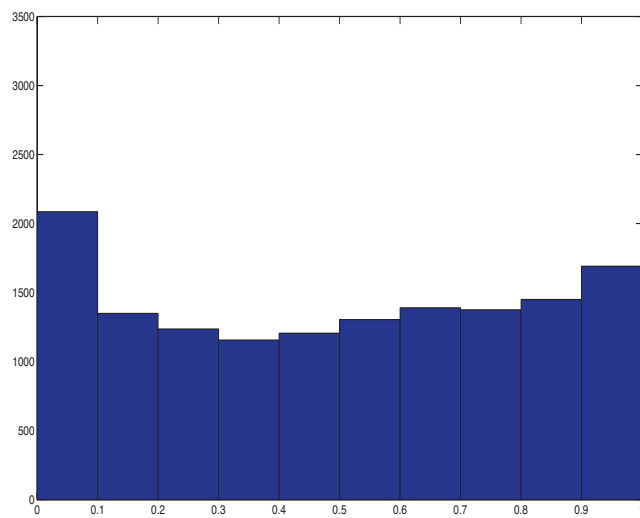
(b) Normal Censurada



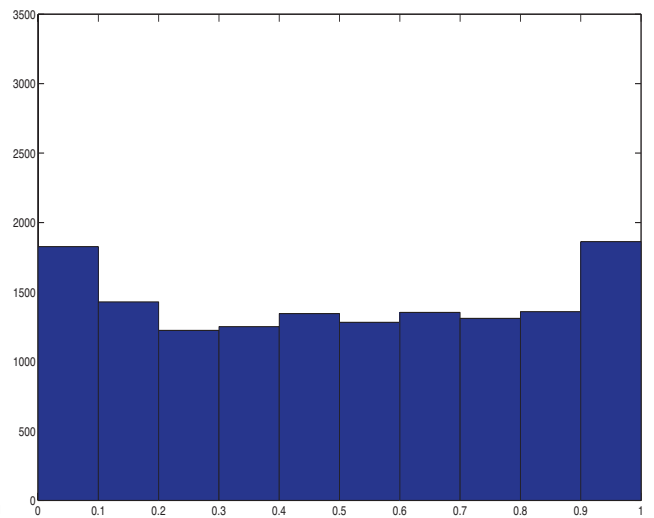
(c) Beta



(d) MEP



(e) Mínimo BS



(f) Mínimo CRPS

Figura 14: Histograma de los valores PIT obtenidos para diferentes densidades predictivas estimadas para un

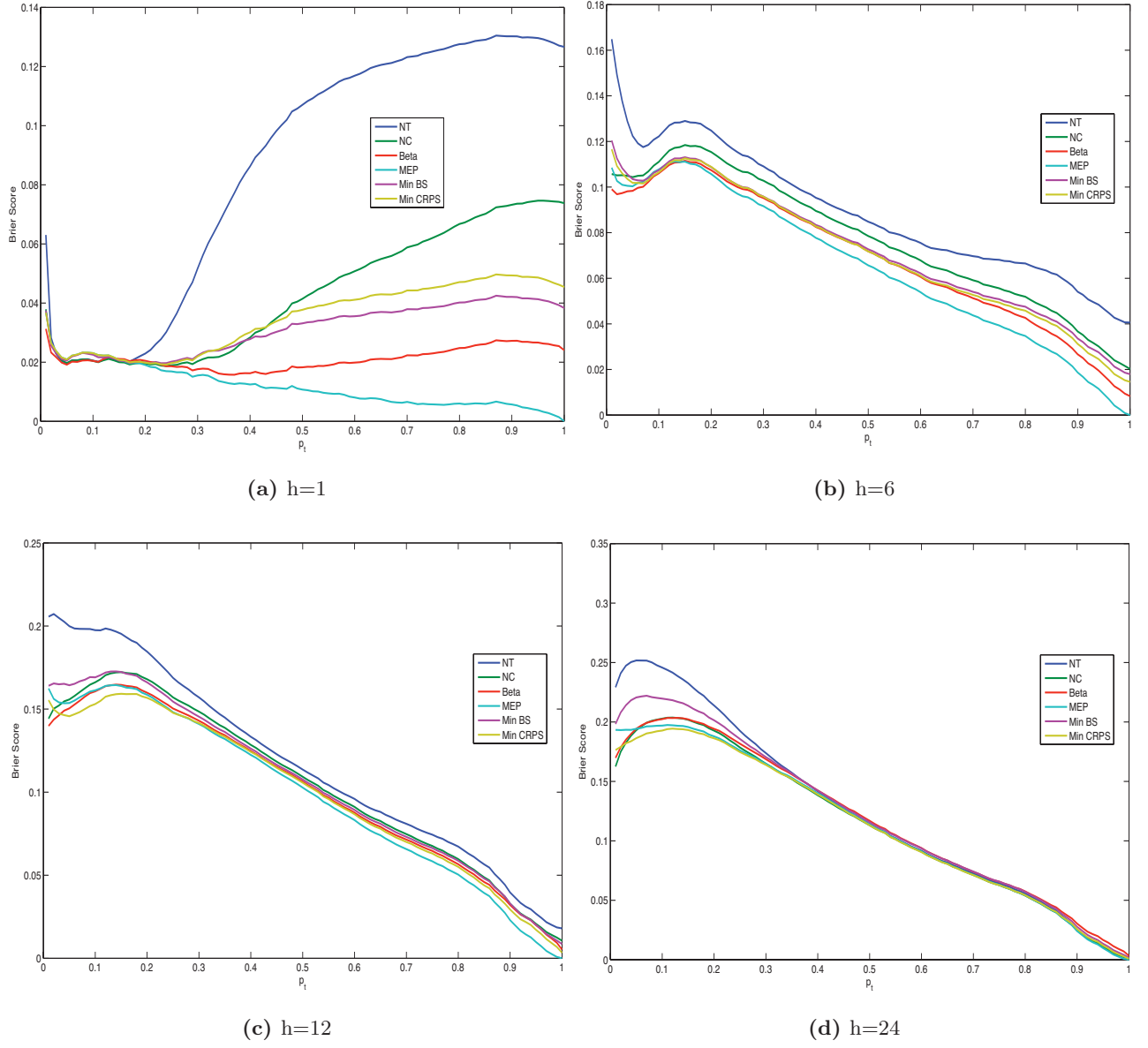
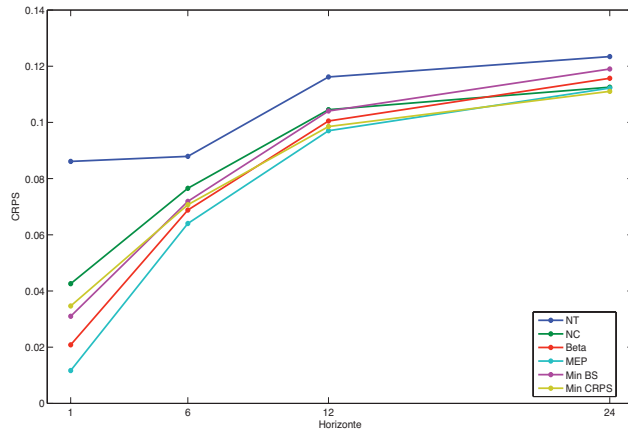
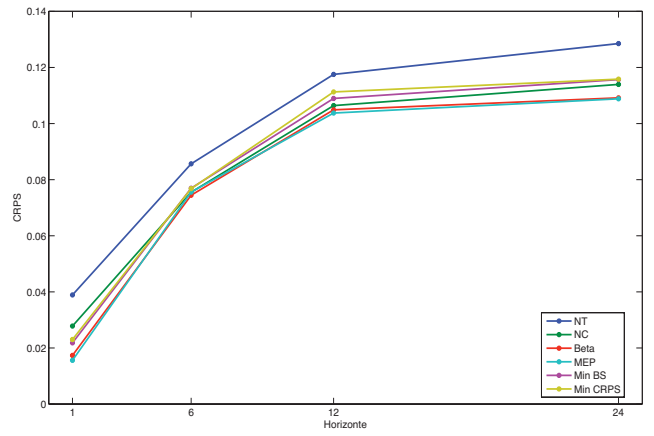


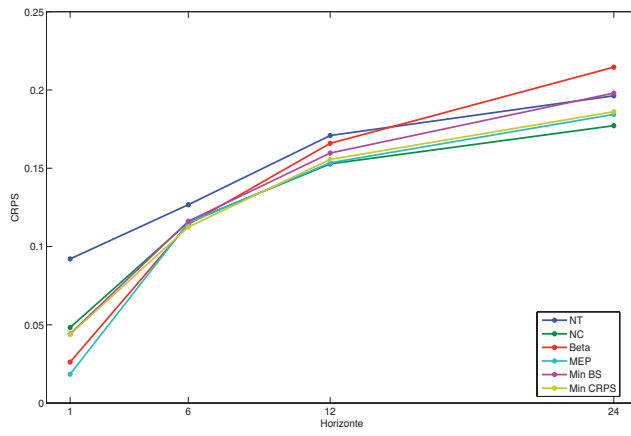
Figura 15: Valor del Brier Score obtenido para las diferentes distribuciones estimadas en un parque eólico empleando para diferentes horizontes de predicción h .



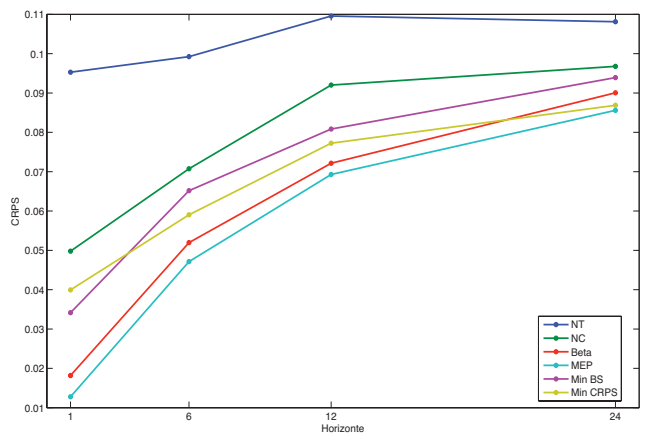
(a) Parque 1



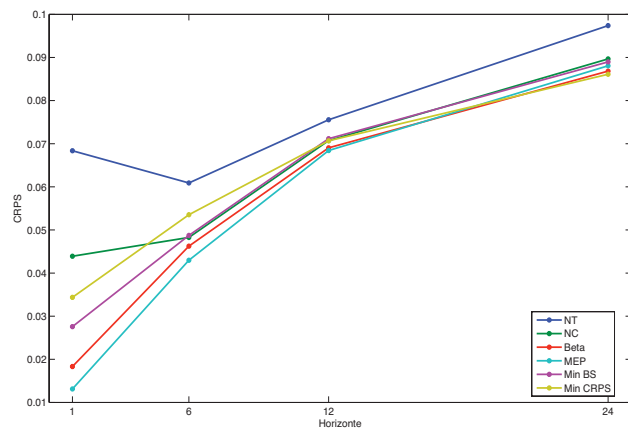
(b) Parque 2



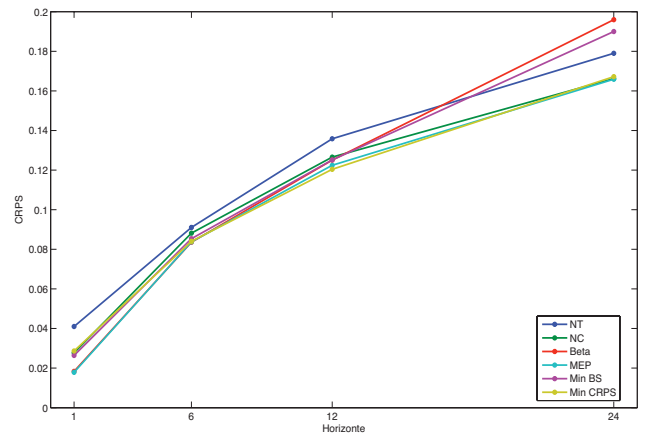
(c) Parque 3



(d) Parque 4



(e) Parque 5



(f) Parque 6

Figura 16: Valor del CRPS para las diferentes distribuciones predictivas empleadas en los diferentes parques eólicos disponibles.

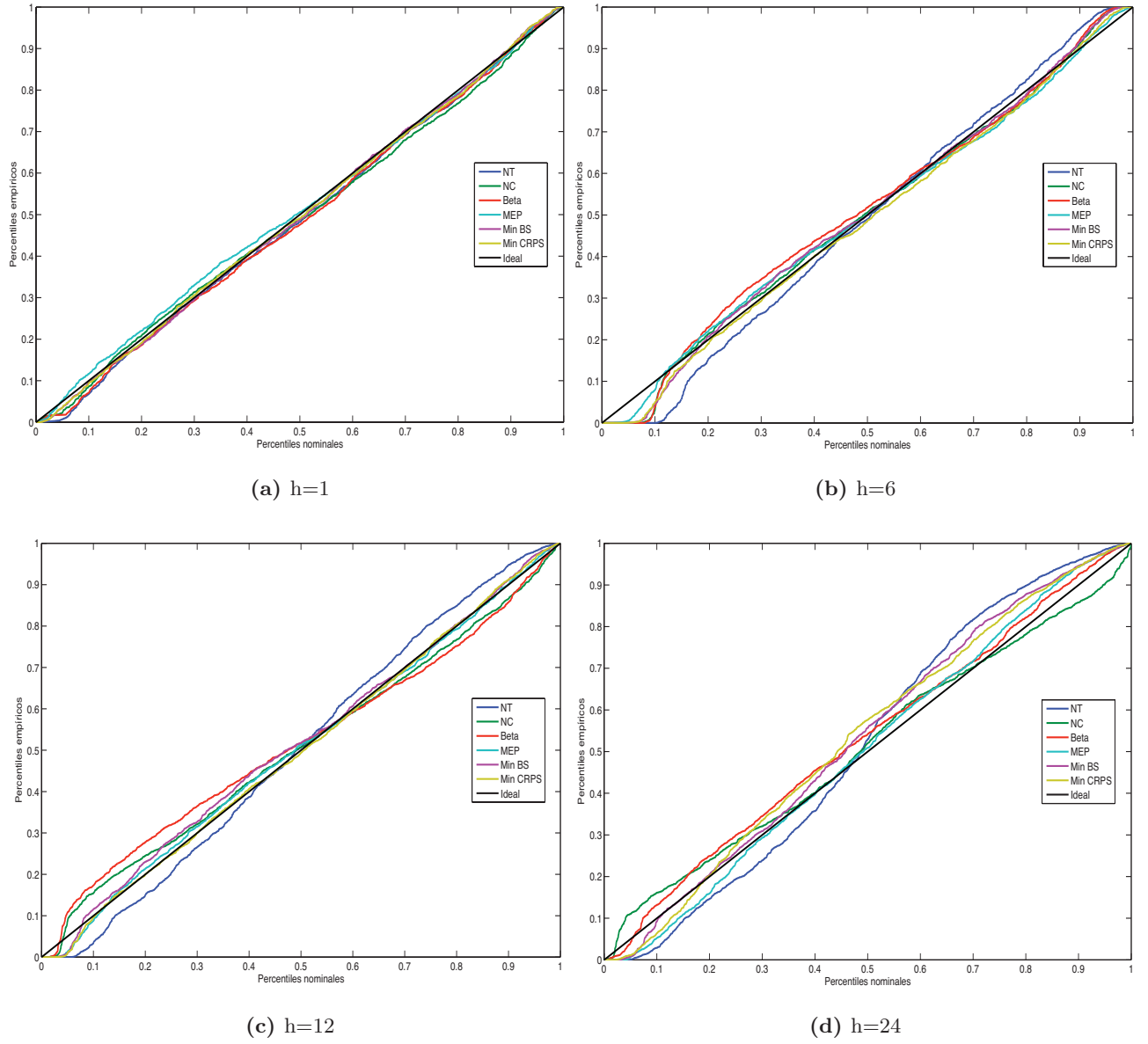


Figura 17: Comparación entre los percentiles obtenidos al evaluar la senda de densidades predictivas estimadas en un parque eólico empleando diferentes distribuciones